CERGE
Center for Economic Research and Graduate Education
Charles University Prague

Learning and Macroeconomic Dynamics

Dmitri Kolyuzhnov

Dissertation

Prague, May 2008

Dmitri Kolyuzhnov

Learning and Macroeconomic Dynamics

Dissertation

Prague, May 2008

DISSERTATION COMMITTEE:

PROFESSOR SERGEY SLOBODYAN, CERGE−EI, CHAIR
PROFESSOR RADIM BOHÁČEK, CERGE−EI
PROFESSOR ANDREAS ORTMANN, CERGE−EI
PROFESSOR PETR ZEMČÍK, CERGE−EI

REFEREES:

PROFESSOR CHRYSSI GIANNITSAROU, UNIVERSITY OF CAMBRIDGE, UK
PROFESSOR KAUSHIK MITRA, UNIVERSITY OF ST. ANDREWS, UK

Abstract

Learning and Macroeconomic Dynamics

by

Dmitri Kolyuzhnov

CERGE–EI, Prague

Professor Sergey Slobodyan, Chair

My dissertation makes a contribution to the field of adaptive learning in macroeconomic models. This contribution is presented in the form of four research papers that constitute different chapters of my thesis.

The first chapter of my dissertation, "Escape Dynamics: A Continuous – Time Approximation" (joint with Anna Bogomolova and Sergey Slobodyan) extends a continuous–time approach to the analysis of escape dynamics in economic models with adaptive learning with constant gain. This approach is based on applying results of the continuous–time version of the large deviations theory to the diffusion approximation of the original discrete–time dynamics under learning. We characterize escape dynamics by analytically deriving the most probable escape point and mean escape time. The continuous–time approach is tested on the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent [61] and Cho, Williams, and Sargent [17].

The second chapter of my dissertation is presented in the paper "Stochastic Gradient versus Recursive Least Squares Learning" (joint with Anna Bogomolova and Sergey Slobodyan), where we perform an in-depth investigation of the relative merits of two adaptive learning algorithms with constant gain, Recursive Least Squares (RLS) and Stochastic Gradient (SG), using the Phelps model of monetary policy studied in the first paper as a testing ground.

The third chapter of my dissertation, "Economic Dynamics Under Heterogeneous Learning: Necessary and Sufficient Conditions for Stability" takes further the issue of different learning of agents, such as RLS and SG learning, in particular the question of stability of equilibrium under the situation when agents differ in the form of adaptive learning algorithms used, in speed of adaptation of their beliefs about the economy to new information, and in initial perceptions, that is the situation of the so-called heterogeneous learning.

The fourth chapter of my dissertation is presented by the paper "Optimal Monetary Policy Rules: The Problem of Stability Under Heterogeneous Learning" (joint with Anna Bogomolova). In this paper we extend the analysis of optimal monetary policy rules in terms of stability of the economy, started by Evans and Honkapohja [31], to the case of heterogeneous learning.

All chapters of my dissertation are related to the question of behavior around the point of equilibrium of the models under adaptive learning of agents. While the first paper analyzes the behavior associated with the tails of the process distribution, that is, escapes out of the point of the self confirming equilibrium (SCE) under homogeneous RLS with constant gain learning, the second paper studies stability of the SCE and escape issues under RLS and SG with constant gain homogeneous learning, comparing dynamics under these two types of learning. The contribution of the third chapter of my dissertation is the derivation of conditions for stability of economic systems under any type of adaptive learning, including homogeneous RLS, homogeneous SG, as well as the general case of heterogeneous mixed RLS/SG learning with different learning adaptation speeds (degrees of inertia) of agents. The fourth chapter uses the stability conditions results derived in the third chapter to extend the analysis of optimal monetary policy rules in terms of stability of the economy, started by Evans and Honkapohja [31], to the case of heterogeneous agents learning.

Professor Sergey Slobodyan
Dissertation Committee Chair

To Anna, Egor and Adelaida

Contents

List of Figures

List of Tables

Acknowledgments

This thesis has grown out of my research (joint with my co-authors, Anna Bogomolova and Sergey Slobodyan) since 2002. During our research we gave presentations, seminars and lectures on the subject of this research at the Central European Program in Economic Theory Workshop, June, 2004, Udine, Italy; Far Eastern Meeting of the Econometric Society, July, 2004, Seoul, Korea; 10th International Conference on Computing in Economics and Finance of the Society for Computational Economics, July, 2004, Amsterdam, the Netherlands; 11th International Conference on Computing in Economics and Finance of the Society for Computational Economics, June, 2005, Washington, D.C., USA; Macroeconomic Workshop, University of Cambridge, November, 2005, Cambridge, UK; European Economics and Finance Society 5th Annual Conference "European Labour Markets, Trade and Financial Flows and Economic Growth", May, 2006, Heraklion, Crete; The congress "50 Years of Econometrics", Econometric Institute, Erasmus University, June, 2006, Rotterdam, the Netherlands; The 12th International Conference on Computing in Economics and Finance Meeting of the Society for Computational Economics, June, 2006, Limassol, Cyprus; European Economic Association and the Econometric Society European Meetings congress, August, 2006, Vienna, Austria; Society for Economic Dynamics Conference, June, 2007, Prague; International College for Economics and Finance of Higher School of Economics, February, 2007, Moscow, Russia; University of Exeter, March, 2007, Exeter, UK; CERGE-EI, Prague, Czech Republic; Novosibirsk State University, Novosibirsk, Russia, and at other events and places. The participants of these events have provided our research with most valuable comments and suggestions. Though it is impossible to list the numerous participants, I express warm thanks to all of them.

However, I would like to specially mention several people whose advise, comments and suggestions have been extremely helpful. I specially mention Sergey Slobodyan, my co-author, who was the first to introduce me to the area of adaptive learning; the members of my dissertation committee: Radim Bohacek, Andreas Ortmann, and Petr Zemcik; Chryssi Giannitsarou and Kaushik Mitra for very helpful refereeing of my thesis and giving valuable suggestions and ideas; Michal Kejak; Jan Kmenta; Kresimir Zigic; and Sarah Peck for English editing of the text of this thesis.

I would like to express special thanks to Seppo Honkapohja who was my supervisor while at Cambridge for drawing my attention to the field of heterogeneous learning and asking to resolve an open question about conditions for stability under heterogeneous learning independent of degrees of inertia of agents.

I gratefully acknowledge the financial support I (and my co-authors) have received from different sources. I express sincere thanks to CERGE-EI, CERGE-EI Foundation and Grant Agency of Charles University for financial support during my study and research, for conferences travel support, for my mobility at the University of Cambridge and at the Economic Development Institute of the World Bank.

I thank my wife and co-author, Anna Bogomolova, and my children, Adelaida and Egor for their help and support, being around and providing me with optimism during the whole process of my research.

Chapter 1

Preface

1.1 Adaptive Learning and its Role in Macroeconomics

Until some time ago, the hypothesis of rational expectations of agents was the major assumption in works studying models of economic dynamics. Later, more and more works appeared that questioned this hypothesis, e.g., Fuchs [35]; Fuchs and Laroque [36]; Grandmont [39, 40]; Grandmont and Laroque [41]; Bray [9]; Bray and Savin [10]; Fourgeaud, Gourieroux, and Pradel [33]; Marcet and Sargent [55]; Evans and Honkapohja [25, 26, 27]; Arifovic [2]; Kirman and Vriend [49]; Cho and Sargent [16]; Marimon [56]; Giannitsarou [37]; Honkapohja and Mitra [43]; Cho and Kasa [15]; and many others. The need to study models under bounded rationality of agents was well argumented in Sargent [60]. Later this approach was also adopted (among others) in the works of Evans and Honkapohja, and a standard argument in defense of bounded rationality can be found in Evans and Honkapohja [29], as well as in Sargent [60].

The rational expectations (RE) approach implies that agents have a lot of knowledge about the economy (e.g., of the model structure and its parameter values), while in empirical work, economists making the RE assumption do not know the parameter values and must estimate them econometrically. According to the argument of Sargent [60], it is more natural to assume that agents face the same limitations economists face and to view agents as econometricians (or statisticians) when forecasting the future state of the economy, who estimate forecasting models using standard statistical procedures, e.g., least squares, and form beliefs about the economic model. The beliefs thus formed are then used to generate agents' actions, and thus influence the realized values of economic variables which are taken as a new data point by the agents. In the next period, agents update their beliefs with the new data. New beliefs then affect actions and economic variables, and this process repeats period after period.

Such a procedure of updating beliefs is called adaptive learning of economic agents and this approach to modeling economic agents constitutes a specific form of bounded rationality and is called the adaptive (econometric) learning[1] approach.

Adaptive learning in macroeconomics plays several roles. First, it is used to test the validity of the rational expectations hypothesis, in particular, whether the dynamics of an economic system with boundedly rational agents may converge to a rational expectations equilibrium (REE) under some learning rule. Second, it is used as a selection device in models with multiple REE. The REE that is chosen is the one that is expected to appear in practice. Third, the dynamics under learning itself could be of interest in a sense that it can be used to explain the behavior of macroeconomic data. And fourth, learning algorithms can be used as method for calculating REE in the model. The advantage of this method is that one can calculate only those REE that could be learned.

The general motivation of my thesis is to make a contribution to the economic literature on adaptive learning described above, to do research in two progressive areas of macroeconomic studies, namely, escape dynamics, which is aimed at exploiting the third role of adaptive learning, and heterogeneous learning, in which adaptive learning appears in its first and second role.

1.2 Escape Dynamics

Adaptive learning of agents generates the fluctuating behavior of macroeconomic indicators, such as inflation and unemployment. These fluctuations can be analytically characterized using two forces that exist in such a dynamic system: mean dynamics, which moves the system to the point of weak convergence, and escape dynamics, which moves the system out.

[1]Note that in my thesis I concentrate on the econometric (statistical) type of adaptive learning. Two other types of adaptive learning approaches: the one that is based on generalized expectation functions (in nonstochastic models) (e.g. papers by Fuchs [35]; Fuchs and Laroque [36]; Grandmont [39, 40]; and Grandmont and Laroque [41]) and the computational intelligence approach, among which are classifier systems, neural networks and genetic algorithms (e.g., papers by Arifovic [2]; Kirman and Vriend [49]; and Cho and Sargent [16]) are described in Evans and Honkapohja [28, pp. 464-465] and are beyond the scope of this thesis.

Major efforts of researchers in this area have been made towards the analytical characterization of the second force, escape dynamics. The approach that has received a bit wider development in the literature, such as in Cho, Williams and Sargent [17] (hereafter, CWS) and Williams [64], is based on the analytical derivation of escape dynamics for the original discrete–time learning process. This approach has both theoretical and practical problems. The theoretical problem is that for cases where a shock to the state process is unbounded (for example, Gaussian), theoretical results allowing full description of escape dynamics are not available. In particular, the theory in this case does not allow one to say what is the most probable direction of deviations from the point of convergence and what is the limiting behavior of the expected time until such an escape.

The practical problem is that the derivation of escape dynamics characteristics for a discrete–time process proposed by Williams [64] implies numerically solving a system of non-linear differential equations with the functions given only numerically. With many lags in the structure of the state process (highly dimensional state variable), this system is hardly solvable.

My work contributes to the current literature on escape dynamics by suggesting another way of deriving analytical characteristics of escape dynamics, which resolves the above mentioned problems. This approach suggests first using a continuous–time approximation of the recursive dynamic system under consideration and then applying the analytical tools developed by Freidlin and Wentzel [34] in order to analytically characterize escape dynamics.

This approach is relatively new in the literature. The only existing paper, Kasa [47], considers this approach only for a very simple one-dimensional case that cannot be easily generalized. Application of the proposed approach can be very useful in complicated non-linear economic models. In case of high dimensionality and lag structure of the underlying model, the costs of deriving escape dynamics for a linearized discrete–time model can be very high compared to the benefits of the continuous–time approach.

The first chapter of my dissertation, "Escape Dynamics: A Continuous–Time Approximation" (joint with Anna Bogomolova and Sergey Slobodyan)[2] extends the continuous–

[2]This paper is published in the CERGE-EI Working Paper Series as working paper number 285. Currently it is on "revise and resubmit" in the *Journal of Economic Theory.* The compressed and abridged four-page version of this paper named "Escape Dynamics: A Continuous Time Approach" is published in the Proceedings of the 4th International Conference on Computational Intelligence in Economics and Finance (CIEF - 2005), Salt Lake City, July 21 -26, 2005.

time approach to the analysis of escape dynamics in economic models with adaptive learning with constant gain.[3] This approach is based on applying the results of the continuous–time version of the large deviations theory to a diffusion approximation of the original discrete–time dynamics under learning. In this paper, we characterize escape dynamics by analytically deriving the most probable escape point and the mean escape time. The proposed continuous–time approach is tested on the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent [61] and CWS. We compare our theoretical results with simulations and the results obtained by CWS.

As a result of our analysis, we express reservations regarding the applicability of the escape dynamics theory to the characterization of the mean escape time for economically plausible values of constant gain in the model of CWS. We show that for these values of the gain, simple considerations and formulae generate much better mean escape time results than the large deviations theory. We explain it by insufficient averaging near the point of the self–confirming equilibrium for relatively large gain values and suggest two changes which might help the approaches based on the large deviation theory work better in this gain interval.

The second chapter of my dissertation is presented in the paper "Stochastic Gradient versus Recursive Least Squares Learning" (joint with Anna Bogomolova and Sergey Slobodyan)[4], where we perform an in-depth investigation of the relative merits of two adaptive learning algorithms with constant gain, Recursive Least Squares (RLS) and Stochastic Gradient (SG), using the Phelps model of monetary policy studied in the first paper as a testing ground. The difference between the two algorithms is that the RLS algorithm[5] has two updating equations: one for updating the parameters entering the

[3] Adaptive learning in such models looks as follows. Each period agents update the parameter estimates in the following way: the updated parameter estimate equals the previous estimate plus some function of the most recent forecast error multiplied by the gain parameter, capturing how important is the forecast error to the agent. In usual recursive least squares derived from OLS, this gain parameter is represented by a decreasing sequence $1/t$. Constant gain learning is used to reflect discounting of the past data by assigning a greater weight to more recent data.

[4] This paper is published in the CERGE-EI Working Paper Series as working paper number 309. Currently it is on "revise and resubmit" in *Macroeconomic Dynamics*.

Note also that though the first two papers have the same co-authors, they are listed in different (not alphabetical) order. It is done in order to highlight the author who wrote the first draft and received the first results of the paper. Thus, the name of this author goes first, then the rest of the authors' names follow in alphabetical order.

[5] The RLS algorithm can be obtained from ordinary least squares (OLS) estimation of parameters, by rewriting it in the recursive form. The generalized RLS is derived from RLS by substituting the gain sequence $1/t$, used in updating the regression coefficients, with any decreasing gain sequence.

forecast functions, the other – for updating the estimates of the second moments matrix of these parameters. The SG algorithm assumes this matrix to be fixed (thus modeling "less sophisticated" agents).

The behavior of the two learning algorithms is very different. Under the mean (averaged) RLS dynamics, the Self–Confirming Equilibrium (SCE) is stable for initial conditions in a very small region around the SCE, and large distance movements of the perceived model parameters from their SCE values, or "escapes," are observed.

On the other hand, the SCE is stable under the SG mean dynamics in a large region. However, the actual behavior of the SG learning algorithm is divergent for a wide range of constant gain parameters, including those that could be justified as economically meaningful. We explain the discrepancy by looking into the structure of eigenvalues and eigenvectors of the mean dynamics map under the SG learning.

The results of our paper suggest that caution is needed when the constant gain learning algorithms are used. If the mean dynamics map is stable but not contracting in every direction and most eigenvalues of the map are close to the unit circle, the constant gain learning algorithm might diverge.

1.3 Heterogeneous Learning

The third chapter of my dissertation, "Economic Dynamics Under Heterogeneous Learning: Necessary and Sufficient Conditions for Stability,"[6] takes further the issue of different types of learning of agents, such as RLS and SG, and in particular the question of stability of an equilibrium when agents differ in a form of adaptive learning algorithms used, in speed of adaptation (degree of inertia) of their beliefs about the economy to new information, and in initial perceptions, that is, the situation characterized as heterogeneous learning.

This paper is devoted to the derivation of necessary conditions and sufficient conditions for stability in the general setup of a stochastic linear forward-looking model of Honkapohja and Mitra [43], which implies possible structural heterogeneity of the model structure and a general case of heterogeneous learning of agents. Studying stability of a dynamic economic model, we may answer the question of a particular economy's behavior

[6]This paper is written as a part of the project "Adaptive learning as a shocks propagation mechanism in the New Keynesian model and optimality and stability of monetary policy rules" supported by the Granting Agency of Charles University, Prague (GAUK), grant number 321/2006/A-EK/CERGE.

around the steady state and also the question of equilibrium selection if the model performs multiple equilibria. However, in models with a very general setup, it may be difficult to find stability conditions that would be computationally tractable and would, additionally, have some economic meaning. This has motivated the research presented in the third chapter.

This paper introduces the concept of δ–stability. Written for heterogeneous learning models, this concept is similar to the E–stability condition — the concept widely used in stability analysis of models with homogeneous learning. This concept follows from the general criterion for stability derived by Honkapohja and Mitra [43] for an economy in the above-mentioned general setup. Their general criterion provides conditions in terms of the model structure and learning heterogeneity, while δ–stability means stability independent of heterogeneity in learning.

In my paper, I derive two groups of sufficient conditions and one group of necessary conditions for δ–stability in terms of the model structure, that is, independent of the type and parametrical characteristics of learning heterogeneity. I have found an easily interpretable unifying condition which is sufficient for convergence of an economy under mixed RLS/SG learning with different degrees of inertia towards a rational expectations equilibrium for a broad class of economic models and a criterion for such a convergence in the univariate case. The conditions are formulated using the concept of a subeconomy and a suitably defined aggregate economy. In the end of the paper I also demonstrate and provide an interpretation of the derived conditions and the criterion on univariate and multivariate examples, including two specifications of the overlapping generations model and the model of simultaneous markets with structural heterogeneity.

The fourth chapter of my dissertation is presented by the paper "Optimal Monetary Policy Rules: The Problem of Stability Under Heterogeneous Learning" (joint with Anna Bogomolova). In this paper we extend the analysis of optimal monetary policy rules in terms of stability of an economy, started by Evans and Honkapohja [31], to the case of heterogeneous learning. Following Giannitsarou [37], we pose the question about the applicability of the representative agent hypothesis to learning. This hypothesis was widely used in the learning literature at early stages to demonstrate convergence of an economic system under adaptive learning of agents to one of the rational expectations equilibria in the economy. We test these types of monetary policy rules defined in Evans and Honkapohja [31] in the general setup of the New Keynesian model that is a work horse of monetary policy models today. It is of interest to see that the results obtained by Evans and Honkapohja [31] for the homogeneous learning case are replicated for the case when the representative agent hypothesis is lifted.

1.4 Summary and Thesis Structure

My dissertation thesis makes several contributions to the field of adaptive learning in macroeconomic models. These contributions are presented in the four research papers that constitute different chapters of my thesis. All chapters of my dissertation are related through the question of behavior around the point of equilibrium of the models under adaptive learning of agents. While the first paper analyzes the behavior associated with the tails of the process distribution, that is, escapes out of the point of the SCE under homogeneous RLS with constant gain learning, the second paper studies stability of the SCE and escape issues under RLS and SG with constant gain homogeneous learning, comparing dynamics under these two types of learning.

The contribution of the third chapter of my dissertation is the derivation of conditions for stability of economic systems under any type of adaptive learning, including homogeneous RLS, homogeneous SG, as well as the general case of heterogeneous mixed RLS/SG learning with different learning adaptation speeds (degrees of inertia) of agents. The fourth chapter uses the stability conditions results derived in the third chapter to extend the analysis of optimal monetary policy rules in terms of stability of an economy, started by Evans and Honkapohja [31], to the case of heterogeneous learning of private agents.

The chapters are divided into two parts reflecting the research area where a particular contribution has been made. Thus the first two papers are collected in Part1 called "Escape Dynamics," while the third and fourth papers are presented in Part 2 that carries the name "Heterogeneous Learning." The Afterword summarizes the results and describes possible directions for further research in the areas of escape dynamics and heterogeneous learning.

Part I

Escape Dynamics

Chapter 2

Escape Dynamics: A Continuous–Time Approximation

Escape Dynamics: A Continuous–Time Approximation*

Dmitri Kolyuzhnov, Anna Bogomolova and Sergey Slobodyan†

CERGE–EI‡
Politických vězňů 7, 111 21 Praha 1,
Czech Republic

Abstract

We extend a continuous–time approach to the analysis of escape dynamics in economic models with constant gain adaptive learning. This approach is based on applying results of continuous–time version of large deviations theory to the diffusion approximation of the original discrete–time dynamics under learning. We characterize escape dynamics by analytically deriving the most probable escape point and mean escape time. The continuous–time approach is tested on the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent [61] and Cho, Williams and Sargent [17] (henceforth, CWS). We compare our results with simulations and the results obtained by CWS. We express reservations regarding the applicability of large deviations theory to characterization of mean escape time for economically plausible values of constant gain in the model of CWS. We show that for these values of the gain, simple considerations and formulae generate much better mean escape time results than the large deviations theory. We explain it by insufficient averaging near the point of self–confirming equilibrium for relatively large gains and suggest two changes which might help the approaches based on large deviation theory to work better in this gain interval.

JEL Classification: C62, C65, D83, E10, E17
Keywords: constant gain adaptive learning, escape dynamics, recursive least squares, large deviations theory

*The earlier version of this paper was published in the CERGE-EI Working Paper Series as w.p. # 285. The current version is on revise and resubmit to the Journal of Economic Theory.

†*{Dmitri.Kolyuzhnov, Anna.Bogomolova, Sergey Slobodyan}@cerge-ei.cz.* *The second author acknowledges support provided by the World Bank through a Ph.D. student fellowship. We thank Klaus Adam, Radim Bohacek, Seppo Honkapohja, Byeongju Jeong, Michihiro Kandori, Michal Kejak, Jan Kmenta, Bruce McGough, Andreas Ortmann, Viatcheslav Vinogradov, and Petr Zemcik, and participants of 2004 CEF, 2004 FEMES, 2004 CEPET, 2005 CEF, and 2006 EEA-ESEM conferences, University of Cambridge macroeconomics workshop, and CERGE-EI seminar for valuable comments and suggestions. All remaining errors are ours.*

‡CERGE–EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic.

2.1 Introduction

The aim of this paper is to extend the continuous–time approach to the analysis of escape dynamics in economic models with adaptive learning and to test it on the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent [61] and Cho, Williams, and Sargent [17]. Incorporating the continuous–time approach into the analysis is motivated by the restricted applicability and computational intensity of the approach used to derive theoretical characteristics of escape dynamics in the recent economic literature. Theoretical analysis of escape dynamics in economic models with adaptive learning allows to theoretically characterize diverse economic phenomena such as currency crises, inflation episodes, endogenous collusion in oligopoly, and cycles of economic activity (see Cho and Kasa [15]; Williams [64, 65, 66, 67]; Bullard and Cho [12]; Cho et al. [17]; Kasa [47]). Escape dynamics also was used to study large mutations in evolutionary games (see Kandori, Mailath, and Rob [45]; and Binmore and Samuelson [5]).

In this literature these phenomena are modeled as a result of escape dynamics in economic models with boundedly rational economic agents who use adaptive learning in the form recently summarized in Evans and Honkapohja [29] to update their beliefs about the economic model. Among the literature devoted to this form of adaptive learning are Bray [9]; Bray and Savin [10]; Fourgeaud, Gourieroux and Pradel [33]; Marcet and Sargent [55]; Evans and Honkapohja [25, 26, 27]; Marimon [56]; and many others. In this literature, agents are considered as econometricians who estimate forecasting models using standard statistical procedures, such as recursive least squares, stochastic gradient, or Bayesian learning, and form beliefs about an economic model. The beliefs thus formed are then used to generate agents' actions, and thus influence the realized values of economic variables, which are taken as a new data point by the agents. In the next period, the agents update their beliefs with the new data. New beliefs then affect actions and economic variables, and this process repeats period after period.

Combined dynamics of parameters describing agents' beliefs and of observed economic variables forms a stochastic recursive algorithm (SRA). Under some regularity conditions, the SRA corresponding to a particular adaptive learning process converges to the rational expectations equilibrium (REE) of the model (or one of the REE in multiple equilibria models), and thus limit dynamics under adaptive learning is the same as that under rational expectations. Stability under adaptive learning which guarantees such convergence has been considered a very important characteristic of the REE in recent monetary policy literature (c.f., Evans and Honkapohja [30] or Bullard and Mitra [11]).

Beyond using the adaptive learning as a *de–facto* equilibrium selection mechanism or a tool for designing policy rules, one could concentrate on the dynamics of the model under adaptive learning as such, in particular, in a case of adaptive learning with a constant gain.[1] In this case convergence of the learning process to REE is only in distribution: there are persistent fluctuations around the REE caused by such learning, and thus rare events — large distance movements called "escapes" — may occur with nonzero probability. During an escape, agents' beliefs about the model move away from nearly rational expectations. As a rule, their actions and the values of realized economic variables also deviate from those observed in the REE.

The analysis of such escape dynamics caused by the adaptive learning process is possible using the theory of large deviations by Freidlin and Wentzell [34] (FW henceforth); Dupuis and Kushner [24]; and others. Depending on what version of the large deviations theory — continuous–time by FW or discrete–time by Dupuis and Kushner [24] — one wants to utilize, there are two possible approaches to the theoretical analysis of escape dynamics: the discrete–time approach and the continuous–time approach. The discrete–time approach, which has received wider attention in the literature, is based on the analytical derivation of escape dynamics for the original discrete–time SRA used to describe a learning process. In the continuous–time approach, a continuous–time diffusion approximation of the discrete–time SRA is derived, and then escape dynamics is studied for this approximation. Kasa [47] applied this approach to a simple one–dimensional model. In this paper, we extend the approach to a multi–dimensional, non–reachable case.

The first approach was used in the majority of the papers cited above, in particular in Cho, Williams, and Sargent [17] (henceforth CWS). These papers work directly with discrete–time SRA processes and use the recent results of Williams [64], who derived numerically the action functional for a linear–quadratic case when the state variable process is autoregressive with Gaussian noise.

There are three basic problems associated with the above approach. First, if the state variable process is subject to unbounded (for example, Gaussian) shocks, the

[1] Constant gain learning discounts the past by assigning more weight to more recent data.

discrete–time version of large deviations theory does not contain theoretical results allowing for a full description of escape dynamics. In particular, the most probable point of escape from the neighborhood of convergence point (as stated above, this point is usually a REE) and the expected time until escape, are unavailable (see CWS, Theorem 5.3). Second, characterizing escape dynamics for the discrete–time process in the way proposed by Williams [64] implies numerical calculation of a functional in a calculus–of–variation problem that leads to a system of non–linear differential equations with numerically derived right–hand side functions. For complicated problems (many lags, high dimensionality), this approach can become numerically intractable. Finally, the analytical solution for escape dynamics of a discrete–time process can be derived only for a restrictive form of learning processes, such as recursive least squares or stochastic gradient learning with a constant gain.

The continuous–time approach developed here resolves these problems. Since diffusion is a natural approximation for a difference equation with Gaussian noise and since FW have developed the theory of large deviations for diffusions, the problem of insufficient theoretical results is removed. The second and the third problems are partially alleviated because the diffusion, derived by approximation around REE — the stationary point of the SRA — is linear. In the large deviations theory, all escape dynamics characteristics — expected time until the beliefs escape any given neighborhood D, the point through which this escape is most likely, and probability of leaving D within a given amount of time — are obtained by minimizing a so–called action functional on the boundary of the neighborhood, ∂D. Given our choice of a linear approximating diffusion, this is a standard linear control theory problem, and the problem of minimizing the action functional is reduced to the trivial task of finding a minimum of a quadratic form on ∂D.

In order to compare the performance of the two approaches of deriving escape dynamics characteristics, the continuous–time approach developed here is tested on the model where the escape dynamics characteristics were already derived using discrete–time approach. This is the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent [61] and CWS.[2]

[2] CWS show that under a constant–gain recursive least squares algorithm, the self–confirming equilibrium (SCE) — a unique set of beliefs corresponding to a time–consistent Nash equilibrium of the RE version of the model — is weakly stable. In this equilibrium, the government believes in strong inflation–unemployment tradeoff. Attempts to exploit this tradeoff, combined with the private sector's rational

The rest of the paper is organized as follows. We briefly describe the dynamic and static versions of the model of CWS in Section 2. We develop the continuous–time approach in Section 3. In Section 4 we present the results of testing the continuous–time approach developed in Section 3 on the model of CWS and compare the approach prediction results with the results of simulations. In Section 5 we discuss the results presented in Section 4 and compare them with the results of CWS, and Section 6 concludes.

2.2 The Model

2.2.1 Setup: Two Versions of the Model

The economy consists of the government and the private sector. The government uses the monetary policy instrument x_n to control inflation rate π_n and attempts to minimize losses from inflation and unemployment U_n. It believes (in general, incorrectly) an exploitable tradeoff between π_n and U_n (the Phillips curve) exists. The true Phillips curve is subject to random shifts and contains this tradeoff only for unexpected inflation shocks. The private sector possesses rational expectations $\widehat{x}_n = x_n$ about the inflation rate, and thus unexpected inflation shocks come only from monetary policy errors.

$$U_n = u - \theta\left(\pi_n - \widehat{x}_n\right) + \sigma_1 W_{1n},\ u > 0,\ \theta > 0, \tag{2.1a}$$

$$\pi_n = x_n + \sigma_2 W_{2n}, \tag{2.1b}$$

$$\widehat{x}_n = x_n, \tag{2.1c}$$

$$U_n = \gamma_1 \pi_n + \gamma_{-1}^T X_{n-1} + \eta_n. \tag{2.1d}$$

Vector $\gamma = \left(\gamma_1, \gamma_{-1}^T\right)^T$ represents government's beliefs about the Phillips curve. W_{1n} and W_{2n} are two uncorrelated Gaussian shocks with zero mean and unit variance. η_n is the Phillips curve shock as perceived by the government, believed to be white noise uncorrelated with regressors π_n and X_{n-1}. Following CWS, we consider two versions of the model: "dynamic" and "static" ones. In the "dynamic" model, vector X_{n-1} contains

expectations, lead to high average inflation. However, inflation periodically performs large deviations, or "escapes," from a neighborhood of the RE Nash equilibrium toward the low inflation time–inconsistent Ramsey outcome of the RE version of the model. This happens when a sequence of stochastic shocks makes the government learn that there is very little tradeoff between unemployment and inflation. These beliefs about the Phillips curve force the government to set inflation low and thus approach the Ramsey outcome.

two lags of inflation and unemployment rates and a constant,

$$X_{n-1} = \left(\begin{array}{ccccc} U_{n-1}, & U_{n-2}, & \pi_{n-1}, & \pi_{n-2}, & 1 \end{array} \right)^T, \tag{2.2}$$

while only the constant is present in X_{n-1} in the "static" version. In other words, the only difference between the two versions of the model lies in the structure of government's beliefs (2.1d), which are more "sophisticated" in the dynamic model. In the sequel we concentrate on the dynamic model and consider the static one only in Section 5.

Given beliefs γ, the government solves

$$\min_{\{x_n\}_{n=0}^{\infty}} E \sum_{n=0}^{\infty} \beta^n \left(U_n^2 + \pi_n^2 \right), \tag{2.3}$$

subject to (2.1b) and (2.1d). This Linear–Quadratic problem produces a linear monetary policy rule

$$x_n = h(\gamma)^T X_{n-1}. \tag{2.4}$$

2.2.2 Nash, Ramsey, and Self–Confirming Equilibria

CWS identify three beliefs consistent with the model. If the government holds Belief 1, $\gamma = \left(-\theta, 0, 0, 0, 0, u(1+\theta^2)\right)^T$, the policy function becomes $x_n = \theta u$. In a model where the government knows the true Phillips curve (2.1a), this is the Nash, or discretionary equilibrium of Sargent [61] and Barro and Gordon [3, 4]. Beliefs 2 of the form $\gamma = (0, 0, 0, 0, 0, u^*)^T$ lead to $x_n = 0$ and zero average inflation for any u^*: Ramsey, or optimal time–inconsistent equilibrium of Kydland and Prescott [53]. Finally, Beliefs 3 where $\gamma_1 + \gamma_4 + \gamma_5 = 0$ (sum of coefficients on current and lagged inflation is zero) asymptotically lead to $x_n = 0$: this is the "induction hypothesis" belief (see Sargent [61]).

In the model with learning, the equilibrium is defined as a vector of beliefs at which the government's assumptions about orthogonality of η_n to the space of regressors are consistent with observations:

$$E\left[\eta_n \cdot (\pi_n, X_{n-1})^T \right] = 0. \tag{2.5}$$

CWS call this point a *self–confirming equilibrium*, or SCE: despite the fact that the government believes into incorrent Phillips curve (2.1d), a particular assumption about it, exemplified by (2.5), turns out to be true. Williams [64] shows that the only SCE in the model is Belief 1.

2.2.3 Adaptive Learning and SRA

In period n, the government uses its current vector of beliefs γ_n to solve (2.3), assuming the beliefs will never change. Thus generated monetary policy action x_n is correctly anticipated by the private sector and produces U_n according to (2.1a). Then the government adjusts its beliefs about the Phillips curve coefficients γ_n and their covariance matrix R_n in an adaptive learning step. Define $\xi_n = \begin{bmatrix} W_{1n} & W_{2n} & X_{n-1}^T \end{bmatrix}^T$, $g(\gamma_n, \xi_n) = \eta_n \cdot \left(\pi_n, X_{n-1}^T\right)^T$, and $M_n(\gamma_n, \xi_n) = \left(\pi_n, X_{n-1}^T\right)^T \cdot \left(\pi_n, X_{n-1}^T\right)$. Next period's beliefs γ_{n+1} and R_{n+1} are given by

$$\gamma_{n+1} = \gamma_n + \epsilon_n R_n^{-1} g(\gamma_n, \xi_n), \tag{2.6a}$$

$$R_{n+1} = R_n + \epsilon_n \left(M_n(\gamma_n, \xi_n) - R_n\right). \tag{2.6b}$$

Equations (2.6) represent a specific form of a recursive learning algorithm. When the gain sequence ϵ_n is given by $1/n$, an appropriate choice of γ_0 and R_0 generates OLS in a recursive form. When $\epsilon_n = const$, this is *constant gain* learning or tracking algorithm.[3]

As $U_n = u - \theta\sigma_2 W_{2n} + \sigma_1 W_{1n}$ and $\pi_n = h(\gamma_n)^T X_{n-1} + \sigma_2 W_{2n}$, the evolution of the state vector ξ_n can be written as

$$\xi_{n+1} = A(\gamma_n)\xi_n + BW_{n+1}, \tag{2.7}$$

where $W_{n+1} = \begin{bmatrix} W_{1n+1} & W_{2n+1} \end{bmatrix}^T$, for some matrices $A(\gamma_n)$ and B. Finally, stack lower–triangular elements of the symmetric matrix R_n into a vector, $vech(R_n)$, and form the parameter vector

$$\theta_n^\epsilon = \begin{bmatrix} \gamma_n^T, & vech^T(R_n) \end{bmatrix}^T \tag{2.8}$$

and the right–hand side vector

$$H(\theta_n^\epsilon, \xi_n) = \begin{bmatrix} \left(R_n^{-1} g(\gamma_n, \xi_n)\right)^T, & vech^T\left(M_n(\gamma_n, \xi_n) - R_n\right) \end{bmatrix}^T. \tag{2.9}$$

Then the dynamics of the model under constant–gain learning can be written as

$$\theta_n^\epsilon = \theta_n^\epsilon + \epsilon H(\theta_n^\epsilon, \xi_n), \tag{2.10a}$$

$$\xi_{n+1} = A(\gamma_n)\xi_n + BW_{n+1}, \tag{2.10b}$$

[3]Constant gain algorithm's assigning more weight to recent data makes sense when agents suspect the world around them to be non–stationary. Presence of sudden breaks in data generating processes, for example as a result of an unpredicted change in the government policy, also calls for tracking algorithms such as constant gain learning. See Evans and Honkapohja [29] for an extensive discussion of constant gain learning and its relation to decreasing gain learning such as OLS.

which is the standard SRA form.[4]

2.3 Continuous–Time Approach

2.3.1 Convergence of SRA and Diffusion Approximation

Define the approximating ordinary differential equations corresponding to our SRA as

$$\begin{aligned} \dot{\gamma} &= R^{-1}\overline{g}(\gamma) = R^{-1}E[g(\gamma,\xi_n)], &\quad (2.11a)\\ \dot{R} &= \overline{M}(\gamma) - R = E[M_n(\gamma,\xi_n)] - R. &\quad (2.11b)\end{aligned}$$

Vector $\overline{\gamma}$ that forms Belief 1 and corresponding 2^{nd} moments matrix $\overline{R}$ are the only equilibrium of the above ODE. This equilibrium is stable. CWS show that under some assumptions, the continuous–time process θ_t^{ε} defined as $\theta_t^{\varepsilon} = \theta_n^{\epsilon}$ for $t \in [n\varepsilon, (n+1)\,\varepsilon)$ converges weakly (in distribution) to $\theta(t,a) = [\gamma^T, vech^T(R)]^T$, solution of the ODE (2.11), where $a = \theta\,(0)$ is the initial condition for the ODE (2.11), and starting point of the process θ_t^{ε}. This solution is also called the "mean dynamics trajectory" of the SRA (2.10), with the right–hand side of (2.11) being the "mean dynamics".

Because of the constant gain learning, the convergence of θ_n^{ϵ} to the mean dynamics trajectory $\theta(t)$ is only weak (in distribution). This implies persistent fluctuations around the trajectory $\theta(t,a)$ and its stationary point $\overline{\theta}$. Large deviations theory studies the probability of rare events, during which these fluctuations force the stochastic process θ_n^{ϵ} out of any given region around the converging trajectory $\theta(t,a)$. Freidlin and Wentzell [34, p.6] state that the probabilities of these rare events "have asymptotics of the form $\exp\left\{-C\varepsilon^{-2}\right\}$ as $\varepsilon \to 0$ (rough asymptotics, i.e., not up to equivalence but logarithmic equivalence)".

The theoretical results of FW on escape dynamics characteristics in continuous time can be applied to the continuous–time approximation of the original discrete–time SRA. This is the essence of the proposed continuous–time approach.[5] Evans and Honkapohja [29, Prop. 7.8] show that as $\varepsilon \to 0$, the process $U_t^{\varepsilon} = \frac{\theta_t^{\varepsilon} - \theta(t,a)}{\sqrt{\varepsilon}}$ converges (weakly) to

[4]Note that vector θ is 27–dimensional, with 6 components representing the government's beliefs γ, and the remaining 21 representing its beliefs about the second moments matrix of γ.

[5]The advantages of the continuous–time approach are discussed in the introduction.

the following diffusion:

$$dU_t^{\epsilon} = D_{\theta}p(\theta\,(t,a))U_t^{\epsilon}dt + \Sigma^{1/2}\,(\theta\,(t,a))\,dW_t, \tag{2.12}$$

where W_t is a multi–dimensional Brownian process with dimensionality equal to that of θ. $p(\theta)$ is the mean dynamics vector, and Σ the matrix whose elements are covariances of different components of the mean dynamics vector, both with respect to the unique invariant probability distribution $\Gamma_{\theta}(dy)$ of the state vector X:[6]

$$p(\theta) \;=\; \int H(\theta, y)\Gamma_{\theta}(dy), \tag{2.13}$$

$$\Sigma_{ij} \;=\; \textstyle\sum_{k=-\infty}^{\infty} Cov\left[H_i(\theta, X_k(\theta)), H_j(\theta, X_0(\theta))\right]. \tag{2.14}$$

This result is used to get continuous–time approximation of SRA, given any initial condition:

$$d\theta_t^{\varepsilon} = D_{\theta}p(\theta\,(t,a))\left[\theta_t^{\varepsilon} - \theta\,(t,a)\right]dt + \sqrt{\varepsilon}\Sigma^{1/2}\,(\theta\,(t,a))\,dW_t. \tag{2.15}$$

Williams [64, Theorem 3.2] shows that the above results can be used to derive a local continuous–time approximation of the SRA around the limit point $\overline{\theta}$ (stable point of the associated ODE (2.11), SCE):

$$d\varphi_t = D_{\theta}p(\bar{\theta})\varphi_t dt + \sqrt{\epsilon}\Sigma^{1/2}(\bar{\theta})dW_t, \tag{2.16}$$

where $\varphi_t = \theta_t - \overline{\theta}$ are deviations from the SCE. The 6×6 upper left corner of $\Sigma(\bar{\theta})$ is equal to the fourth moments matrix Q of CWS evaluated at $\overline{\theta}$. Matrices $D_{\theta}p(\bar{\theta})$ and $\Sigma(\bar{\theta})$ need to be evaluated only at the SCE. This could be performed analytically (the technical appendix with these derivations is available from the authors upon request).

Diffusion (2.16), used in this paper, approximates a highly nonlinear multidimensional SRA only at the stationary point of the mean dynamics. Dembo and Zeitouni [22, p. 223] argue that "the rationale here is that any excursion off the stable point has an overwhelmingly high probability of being pulled back there, and it is not the time spent near any part of ∂D that matters but the *a priori* chance for a direct, fast exit due to a rare segment in the Brownian motion's path."

[6] State vector ξ_n has a unique invariant probability distribution: it contains stationary Gaussian random variables W_{1n} and W_{2n}, a constant, and a stable 4–dimensional AR(1) variable. This distribution can be calculated explicitly.

2.3.2 Action Functional and Escapes

Suppose that we have a stochastic process, for example, some diffusion. The basic idea of the theory of large deviations for paths of stochastic processes is that the probability of stochastic process's deviating from a given path along a specific trajectory can be determined by the value of a certain functional (called *action functional*) on this trajectory. Action functional $I_{0T}(\varphi)$ represents the costs associated with moving along some trajectory φ for a period of time $[0, T]$. Cost function $I(T, x, y) = \min_{\varphi_0 = x, \varphi_t = y} I_{0t}(\varphi)$ is the minimal cost required for transition from x to y in time T. *Quasipotential* $I(x, y) = \inf_{T>0} I(T, x, y)$ is the minimal cost necessary to move from x to y given arbitrary (potentially infinite) time. The idea here is that the system moves in the direction along which it incurs the least cost.

Suppose that such a functional exists. We are given some neighborhood D of a stationary point of the diffusion's drift, O. Under certain assumptions one can derive the probability that a stochastic process belongs to D from the minimum value of the quasipotential $I(O, y)$ on the boundary of D, $\{y : y \in \partial D\}$. The most probable point at which the stochastic process leaves (escapes) D, is the point where $I(O, y)$ has a minimum. The minimum of $I(O, y)$ also allows one to derive asymptotic behavior of the mean escape time, *i.e.*, the expected time needed for the stochastic process to cross the boundary of D for the first time.

The exact results on the mean exit time and the dominant escape point are given in Dembo and Zeitouni [22, Theorem 5.7.11]. In particular, the limiting behavior of the mean escape time, $E_x(\tau^{\varepsilon})$, is given by

$$\lim_{\epsilon \to 0} \varepsilon \ln E_x(\tau^{\varepsilon}) = \overline{I}, \tag{2.17}$$

where $\bar{I}$ is a minimum value of the quasipotential on ∂D. The most probable escape point is an extremal of the quasipotential on ∂D; see Appendix A.1 for exact definitions.

2.3.3 Minimizing the Action Functional

For a diffusion $d\varphi_t = A\varphi_t dt + \sqrt{\epsilon} B dW_t$, Dembo and Zeitouni [22, p. 214] provide the following expression for the action functional:

$$I_{0T}(\varphi) = \inf \frac{1}{2}\int_0^T \left|\dot{g}_t\right|^2 dt, \tag{2.18a}$$

$$\text{s.t. } \dot{\varphi}_t = A\varphi_t + B\dot{g}_t, \tag{2.18b}$$

$$\varphi_0 = 0, \tag{2.18c}$$

where a stationary point of the drift O is assumed to be the origin. Minimization is performed over all possible trajectories of $\dot{g}_t = u_t$, which take the system from the origin to φ_T in exactly T time units. In the approximating diffusion (2.16) the matrix A equals $D_\theta p(\bar{\theta})$, and $B = \Sigma^{1/2}(\bar{\theta})$.

The only complication with this formulation stems from the fact that matrix $B = \Sigma^{1/2}(\bar{\theta})$ can be singular.[7] As a result, there might be points in the state space that could not be reached in any time using any control trajectory $\{u_t\}_{t=0}^{\infty}$: the system (A, B) is not necessarily reachable.[8] The way to proceed with the control problem for an unreachable system is to transform the state space so that first k new coordinates (z_1) form the basis of the *reachable subspace*, where the remaining $n-k$ (z_2) coordinates all equal zero. In these coordinates, the system's evolution on the reachable subspace is governed by

$$\dot{z}_1 = \overline{A}_1 z_1 + \overline{B}_1 u, \tag{2.19}$$

where $z_1 = (T_1)^T \varphi$, T_1 is the basis of the reachable subspace, and the system $(\overline{A}_1, \overline{B}_1)$ is by construction reachable; see Dahleh et al. [21, Ch. 22] for the construction. The action functional (2.18) is then rewritten as

$$I_{0T}(z_1) = \inf \frac{1}{2}\int_0^T \left|\dot{g}_t\right|^2 dt, \tag{2.20a}$$

$$\text{s.t. } \dot{z}_1 = \overline{A}_1 z_1 + \overline{B}_1 \dot{g}_t, \tag{2.20b}$$

$$z_1(0) = 0. \tag{2.20c}$$

[7] It turns out that it is singular in the model of CWS. Singularity comes from collinearity of regressors at the SCE: inflation rate and unemployment rate equal a constant plus *i.i.d.* noise. As a result, 14 out of 21 entries in $\overline{R}$ are constants which do not depend on the noise magnitude, and the rank of matrix Σ equals $27 - 14 = 13$.

[8] A general way of dealing with singular diffusion matrices is to consider two subsystems, such that the difusion matrix is uniformly positive definite in one of them, and no diffusion is present in the second subsystem; see Roy [59].

To find $\bar{I}$, one has to minimize $I_{0T}(z_1)$ over the time to escape T and all points $z_{1,D}$ such that $T_1 z_{1,D} \in \partial D$. In other words, the problem of finding the minimum value of the action functional over all trajectories starting at the origin and terminating on ∂D in an arbitrary time is split into two separate problems: first, find the minimum norm control path, $\dot{g}_t$, which takes the linear control system from the origin to $z_{1,D}$ in arbitrary time, and then minimize over all possible terminal points $z_{1,D}$.

The first problem is a standard control problem that has the following solution:

$$I(z_{1,D}) = \frac{1}{2} z_{1,D}^T \cdot \overline{G}^{-1} \cdot z_{1,D}, \tag{2.21}$$

where $\overline{G}$ is Gramian of the reachable subsystem. See Appendix A.2 for details and definitions of matrices T_1 and $\overline{G}$. The problem of finding the minimum value of the action functional then becomes a trivial one: minimize the quadratic function of $z_{1,D}$ on $\{z_{1,D} \colon T_1 z_{1,D} \in \partial D\}$. By solving this problem, we find the most probable point of escape, $T_1 z_{1,D}$, and the rate of convergence, $\overline{I}$, that characterizes the limiting behavior of the mean escape time by the limit expression $\lim_{\epsilon \to 0} \varepsilon \ln E_x(\tau^\varepsilon) = \overline{I}$.

2.4 Testing the Approach on the Phelps Problem

2.4.1 Simulations and Reduced Dimensionality of the Model

In this section we present the results of applying the continuous–time approach developed in the previous section to the Phelps problem studied previously by CWS, who exploited the discrete–time approach. All the model parameters are taken to be the same as in CWS in order to guarantee the possibility of comparison between the results.[9] Before applying the theory, we first run simulations of the model to have an initial picture of the dynamics in the system.

Following CWS, we plot simulation runs in two–dimensional space, where the abscissa is set to be the "inflation slope coefficient" (the sum of beliefs coefficients before inflation and lagged inflation, $\gamma_1 + \gamma_4 + \gamma_5$) in order to see how the system moves towards the "induction hypothesis" plane, $\gamma_1 + \gamma_4 + \gamma_5 = 0$. In contrast to CWS, we use as an ordinate the intercept coefficient summed with the "lagged unemployment slope coefficient" multiplied by the average unemployment rate, $\gamma_6 + u \cdot (\gamma_2 + \gamma_3)$, rather than the intercept

[9]Throughout the paper we use the same parameter values as in CWS: $\sigma_1 = \sigma_2 = 0.3$, $u = 5$, $\theta = 1$, $\beta = 0.98$. All the figures are for simulations with $\epsilon = 0.001$, unless otherwise noted.

coefficient, γ_6, alone. The exact algebraic form of the coordinates used is explained by the following consideration.

Suppose that the government's beliefs are fixed for a number of periods at γ_n, so that the state dynamics becomes stationary, with an unconditional expectation of U_n being u and that of π_n being some $\widetilde{\pi}$. What is the expected value of η_n? $\eta_n = U_n - \gamma_1\pi_n - \gamma_{-1}^T X_{n-1}$, and so

$$E[\eta_n] = u - (\gamma_1 + \gamma_4 + \gamma_5)\widetilde{\pi} - (\gamma_2 + \gamma_3)u - \gamma_6. \tag{2.22}$$

This back–of–the–envelope calculation suggests that from the government's point of view, linear combinations $\gamma_1+\gamma_4+\gamma_5$ and $(\gamma_2+\gamma_3)u+\gamma_6$ rather than the whole vector γ matter. As it is exactly a perceived error η_n which matters for the adjustment of θ in (2.6), one presumes that coordinates

$$(\widetilde{\gamma}_1, \widetilde{\gamma}_2) = \Big(\ \gamma_1 + \gamma_4 + \gamma_5, \quad u \cdot (\gamma_2 + \gamma_3) + \gamma_6\ \Big) \tag{2.23}$$

are useful in thinking about the model.

The above coordinates are used to plot a typical simulation run started at SCE with $\epsilon = 0.001$, including an escape towards the "induction hypothesis" belief (Belief 3 of CWS) and very low inflation. In the $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ plane, all simulation points are very close to a one–dimensional curve (a straight line). Very similar graphs are obtained in all 1000 runs, which strongly suggests that we could use coordinates $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ to effectively reduce the dimensionality of the system.

The reason for such almost "one–dimensionality" of the dynamics can be found by looking at the parametrical structure of the SRA. Consider equation (2.6a). At the SCE,[10] the largest eigenvalue of $\overline{R}^{-1}$ is λ_1 =3083.8 and the next largest λ_2 =29.1, less that 1% of λ_1.[11] Therefore, if one writes $g(\gamma_n, \xi_n)$ as a linear combination of eigenvectors of $\overline{R}^{-1}$, then the projection of $g(\gamma_n, \xi_n)$ onto v_1, the eigenvector corresponding to λ_1, is magnified 100 times as strongly as the projection onto v_2. The dynamics described by $\overline{R}^{-1} g(\gamma_n, \xi_n)$ is, thus, almost 1–dimensional. In coordinates $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$, v_1 is approximately proportional to (1,-5)´. Figure 2.1 shows that, indeed, all the simulation run points are aligned along this vector.

[10] SCE is a starting point of the process for all simulations and theoretical derivations.

[11] This dramatic difference is due to the fact that u is so large. Entries of $\overline{R}$ are of order u^2, u, and 1. Large u leads to significantly different entries of $\overline{R}$ and thus to a large λ_1/λ_2 ratio. For $u = 1$, the ratio λ_1/λ_2 drops to 5.17.

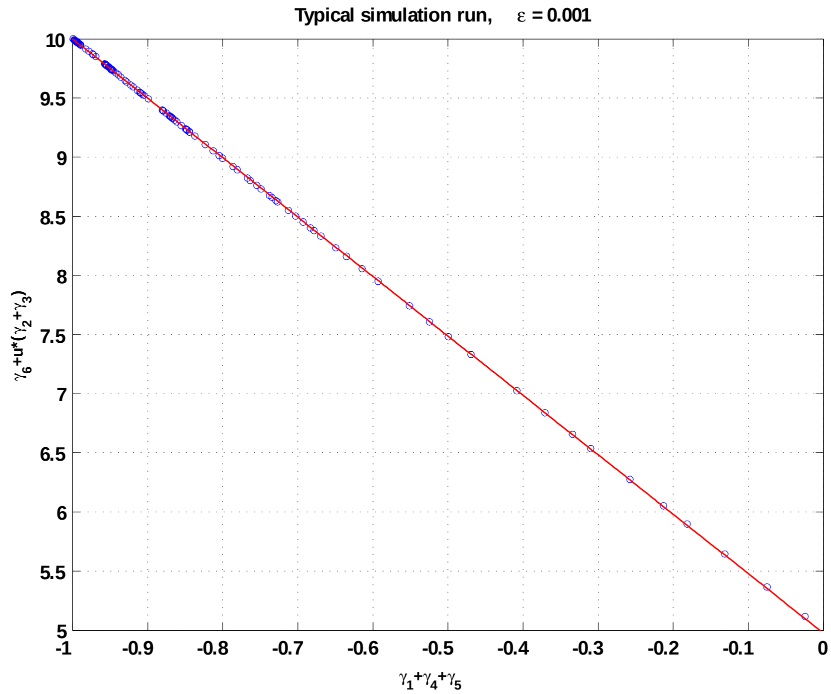

Figure 2.1: Typical simulation run and the "largest" eigenvector of R^{-1}.

One cannot help noticing the striking similarity between our Figure 2.1 and Figure 6 of CWS, which uses coordinates $(\widehat{\gamma}_1, \widehat{\gamma}_2)=\Big(\gamma_1 + \gamma_4 + \gamma_5, \quad \gamma_6 \Big)$. In $(\widehat{\gamma}_1, \widehat{\gamma}_2)$ space v_1 is proportional to (1,-7.86)´, which is again sufficiently close to the line drawn by simulation run points of CWS. Therefore, straight lines drawn in different coordinates by CWS and by us in this paper are nothing but projections of the "largest" eigenvector of $\overline{R}^{-1}$, v_1, onto different hyperplanes.[12]

It is possible to neglect the dynamics of R in considering escapes because the covariance matrix $\Sigma_{\overline{\theta}} = cov[H(\overline{\theta}, X(\overline{\theta}))]$ contains $\overline{R}^{-1}$ in its upper–left corner, and its largest eigenvalue's eigenvector $\widetilde{v}_1$ is proportional to (1,-5)´ in $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ coordinates. The parametrical structure of $\Sigma_{\overline{\theta}}$ is close to block–diagonal, so there is very little interaction between RHS terms of (2.11), $R_n^{-1} g(\gamma_n, \xi_n)$ and $M_n(\gamma_n, \xi_n) - R_n$, which influence components of γ and of R, at least for "largest" eigenvectors which determine the dynamics

[12] The fact that simulation run points plot almost an ideal straight line suggests that the matrix R does not change much along the typical escape path, preserving the ratio of the largest to the second largest eigenvalue and the direction of the "largest" eigenvector. This conjecture turns out to be correct for relatively large values of ϵ.

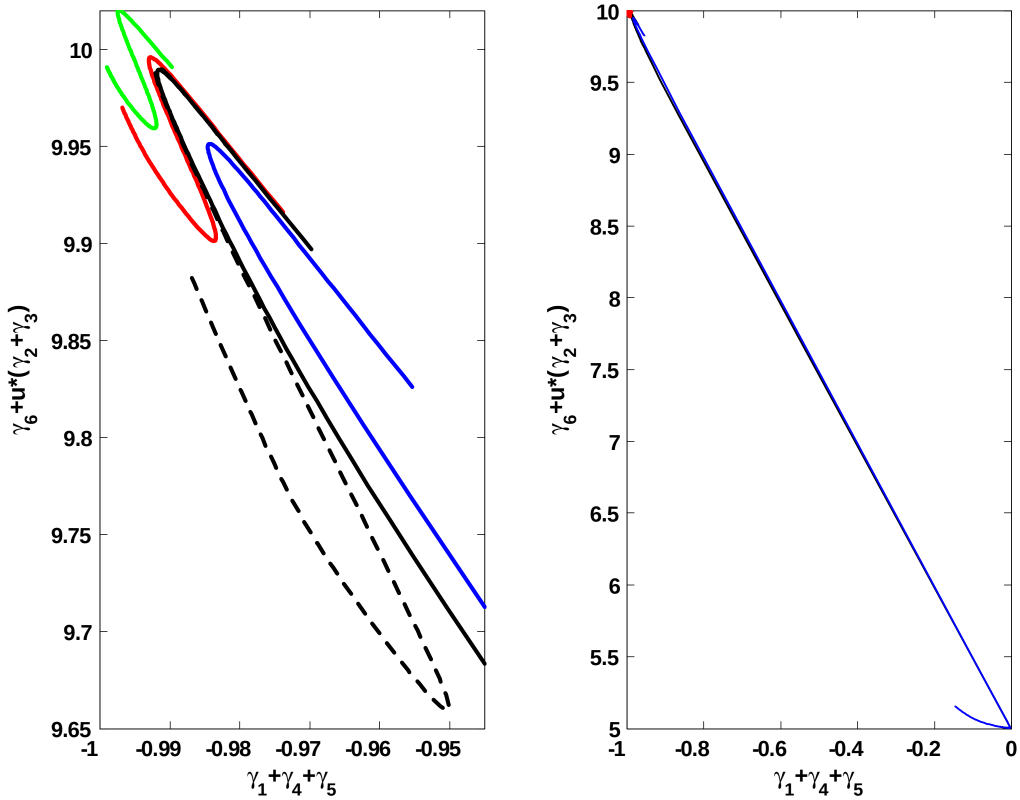

Figure 2.2: The mean dynamics trajectories.

of the model.[13]

After reducing the dimensionality of the model, it is possible to analyze the behavior of the mean and the stochastic part of dynamics using the simulations results. Several trajectories of the mean dynamics of the model, given by (2.11), are presented in Figure 2.2 in $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ coordinates. The region around the SCE where the mean dynamics points back towards it is very small; if the initial deviation from the SCE is relatively large, the mean dynamics trajectory treks towards Belief 3, or the "induction hypothesis" plane, where $\widetilde{\gamma}_1$=0. After spending some time in the neighborhood of $\widetilde{\gamma}_1$=0, the trajectory slowly returns back to the SCE. The right panel of Figure 2.2 tracks several mean dynamics trajectories as they travel to the "induction hypothesis" plane. The paths are almost

[13]Matrix $\Sigma_{\overline{\theta}}$ is close to being block–diagonal, similarly to $R(\overline{\theta})$ in Evans and Honkapohja [29, Eq. 14.6], in the following sense: If one takes one "largest" eigenvector of 6×6 upper left corner of $\Sigma_{\overline{\theta}}$ and two "largest" eigenvectors of 21×21 lower right corner and pads them with zeros appropriately, the resulting vectors are almost indistinguishable from the three largest eigenvectors of the whole $\Sigma_{\overline{\theta}}$ (for a block–diagonal matrix, they should be exactly equal).

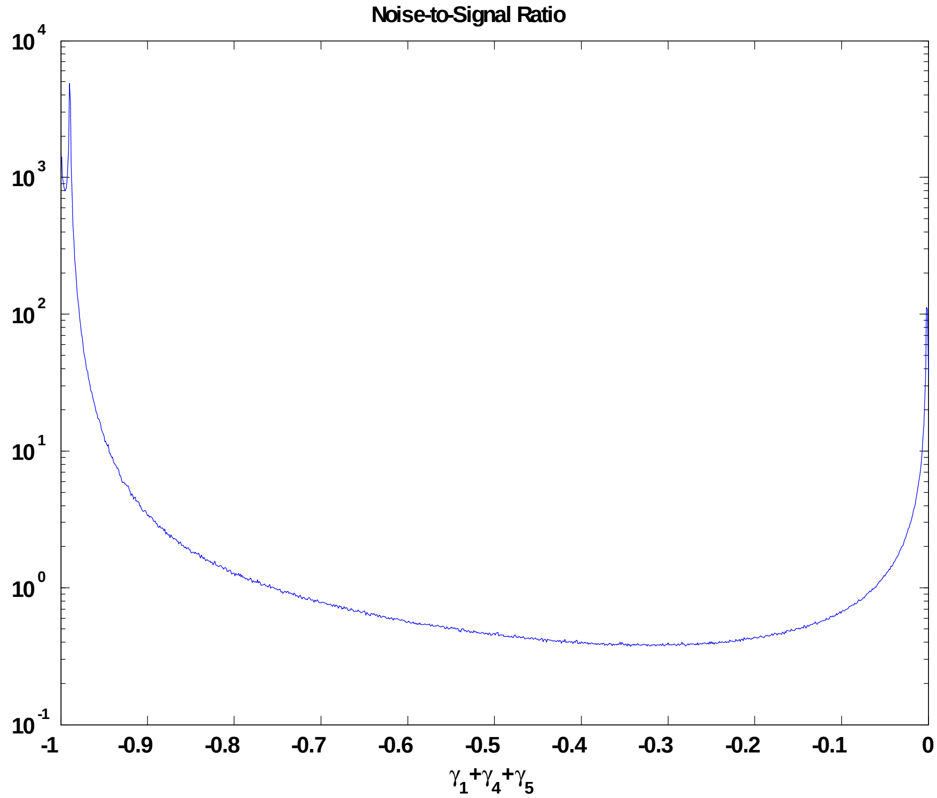

Figure 2.3: Noise-to-signal ratio along the "largest" eigenvector of R^{-1}.

indistinguishable at this scale: Away from the immediate neighborhood of the SCE, mean dynamics trajectories are rapidly converging to the line connecting the SCE with (0,5).

To understand the relation between the mean and the stochastic parts of the dynamics of (2.6), consider Figure 2.3. It plots a ratio of the relative magnitude of the stochastic dynamics, given by $R_n^{-1}\left\{g(\gamma_n,\xi_n)-E[g(\gamma_n,\xi_n)]\right\}$ averaged over 4000 realizations of ξ_n, and of the mean dynamics $R_n^{-1}E[g(\gamma_n,\xi_n)]$. The ratio is evaluated at different points θ along the eigenvector $\widetilde{v}_1$. For large deviations from the SCE the mean dynamics dominates the stochastic part. In the small region around SCE where the mean dynamics points back towards it, stochastic dynamics is on average hundreds and thousands times larger than the mean dynamics.[14]

[14]To understand the importance of this point, consider the heuristic derivation of the mean dynamics ODE in Evans and Honkapohja [29, p.127]. N steps ahead value of θ, θ_{n+N}, is given approximately as

$$\theta_{n+N}\approx\theta_n+(N\gamma)\,\frac{1}{N}\sum_{i=0}^{N-1}H\left(\theta_n,X_{n+1+i}\right)\approx\theta_n+N\gamma h\left(\theta_n\right).$$

The last approximation is justified by invoking the law of large numbers. But as we have shown in Figure

Using the theoretical derivations and the knowledge of the system obtained through simulations, we think about escapes in the following way. Consider a small neighborhood D of the SCE. After the trajectory crosses the boundary ∂D, we assume that the stochastic dynamics does not play any role, and the model's behavior is determined exclusively by its mean dynamics (2.11). Arguments presented in the previous paragraph allow us to claim that this is a very good approximation far from the SCE. Moreover, as all mean dynamics trajectories are very close in this region, one does not need to know the exact escape point to predict the most likely behavior of the system during travel to the low inflation outcome. A process of excursion towards $\widetilde{\gamma}_1 = 0$ is, therefore, split into two parts: first, stochastic "escape" from D, and second, almost deterministic movement to $\widetilde{\gamma}_1 = 0$ and back to the SCE. If our selection of ∂D is such that after crossing it the mean dynamics points towards $\widetilde{\gamma}_1 = 0$, there is no additional contribution to the quasipotential, as the system does not need any additional energy to move away from the SCE. Therefore, we concentrate on the first part, the stochastic "escape" from the set D, such that after the escape the mean dynamics points away from the SCE.

2.4.2 Analytical Results vs. Simulations: Point of Escape

We select the set D and its boundary ∂D in several ways. The first way is to use a cylinder: a sphere in six–dimensional γ space, and no binding restrictions in 21–dimensional space of components of R. This approach is similar to the road taken by CWS. The theoretical results of the problem of minimizing action functional on the cylinder are presented in Appendix A.2. This simple calculation does not replicate the behavior observed in simulations for gain value $\epsilon = 0.001$:[15] no matter what the cylinder's radius is, the escape from the SCE is predicted to occur in the approximate direction (1,-6.5)´ in $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ coordinates. When one selects the cylinder's radius so that the cylinder crosses $\widetilde{v}_1$ at $\widetilde{\gamma}_1$=-0.985 (more on this choice below), the distance between the mean of observed escape points and the theoretically derived escape point, $d1$, equals 100% of the distance from the SCE to the theoretical escape point. Figure 2.4 presents a histogram of observed

2.3, the mean value of $|H(\theta_n, X_{n+1+i}) - h(\theta_n)|$ is overwhelmingly large relative to the $h(\theta_n)$. Under these circumstances, $\sum_{i=0}^{N-1} H(\theta_n, X_{n+1+i})$ could be very different from $h(\theta_n)$ unless N is very large.

[15] Gain value $\epsilon = 0.001$ is chosen for a comparison of simulations with analytical results as the lower boundary for economically plausible values of ϵ. Large deviations theory has to work for 0.001, if it works for economically plausible values of $\epsilon \in [0.001, 0.01]$ at all.

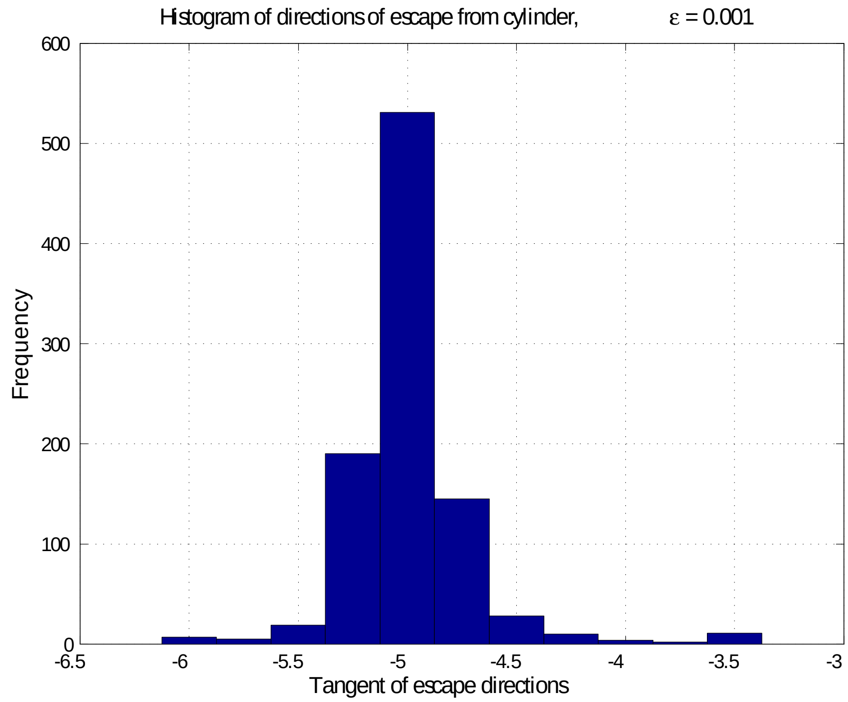

Figure 2.4: Histogram of tangents of escape directions in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space for $\epsilon = 1 \cdot 10^{-3}$. Cylinder with the radius described in the text.

escape directions.[16] As is easy to see, the theoretical prediction of approximately -6.5 is way off the mode of the empirical distribution, which is -5. Fewer than 1% of simulation runs result in escape in the direction with tangent less than -6. Therefore, we conclude that the cylinder is not a good choice for the escape region D, at least at this value of ϵ. Additionally, simulation runs show that many of the "escapes" generated in this way violate our basic assumption: mean dynamics paths, initiated at the escape point, do not deviate towards Belief 3, see Figure 2.5, where such points are marked by green crosses.

Our second way of selecting the boundary of the set D is the curve in $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ space, defined numerically as a path "separating" trajectories coming back to the SCE from those which first travel to the "induction hypothesis" plane under the mean dynamics.[17] We derive this curve as follows. Find two points on the eigenvector $\widetilde{v}_1$, such that one

[16]These tangents are obtained as follows. For a given simulation run, determine the point of escape from the cylinder and project it into $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ space. Write the vector which starts at the SCE and points towards the escape point as $(1, Tg)'$. Then Tg is simply a tangent of the angle between $(1,0)'$ and $(1, Tg)'$. Figure 2.4 shows the histogram of Tg.

[17]This surface is not a separatrix in the strict sense of the word, as all trajectories eventually return back to the SCE and asymptotically converge to it. However, there is a sensitive dependence on the initial conditions in the neighborhood of this surface.

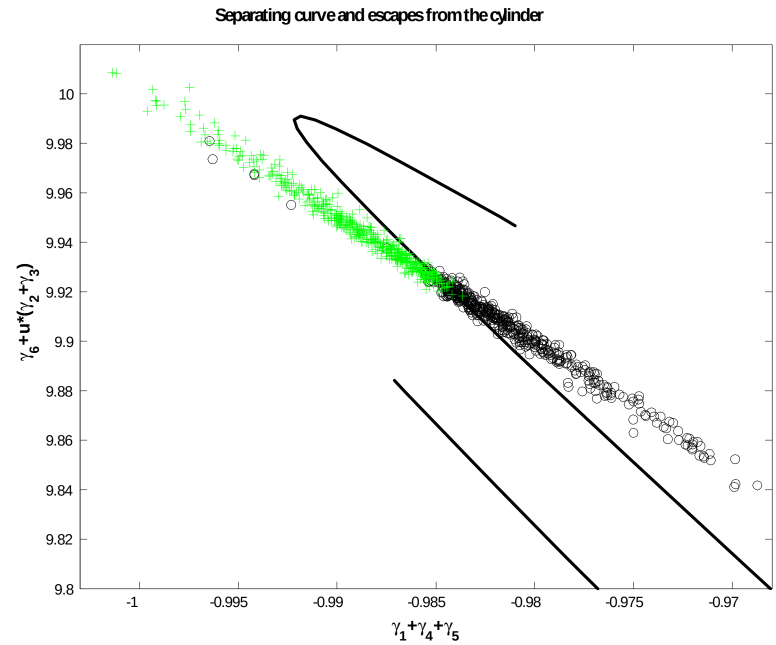

Figure 2.5: Continuation of escape runs under mean dynamics. Runs that ended in a point marked by '+' return to the SCE immediately, those marked by 'o' come back only after an excursion to the "induction plane".

of them generates a trajectory of (2.11) immediately coming back to the SCE, while another starts an excursion toward the induction hypothesis plane. Using a sequence of binary bisections, shrink the interval between two such points to any desired small number. Starting from this point, solve (2.11) forward and backward in time. Project the resulting trajectory into $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ space and call this curve ∂D. In the left panel of Figure 2.2, we plot two trajectories that turn back to the SCE (two leftmost lines), and two paths travelling to the "induction hypothesis" plane first and then coming back (two rightmost lines). The "separating" curve is plotted as a dashed line. At the point where the "separating" curve intersects with eigenvector $\widetilde{v}_1$, $\widetilde{\gamma}_1$ approximately equals -0.985. Radius of the cylinder, described above, was selected in such a way that it intersects $\widetilde{v}_1$ at the same point as the "separating" curve.

Deriving a "separating" surface has the advantage of taking into account some information about behavior of the mean dynamics away from the SCE. The linear approximating diffusion (2.16) discards this information by taking into account only $D_\theta p(\bar{\theta})$ rather than $D_\theta p(\theta)$. On the other hand, this procedure is very simplistic and heuristic: assuming that such a separating surface exists in the original 27–dimensional space, there is no particular reason to believe that its projection into two–dimensional $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ space coincides with, or is concentrated around, the projection of one particular path. Nevertheless, Figure 2.5 shows that empirically the separating curve does a good job at least in some respects: consider projections of all 1000 simulated points of escape from the cylinder described above. We use these escape points as initial values for the trajectory of (2.11). If this trajectory comes back to the SCE, projection of the escape point is plotted using a '+' symbol in Figure 2.5. If the path first travels to the "induction hypothesis" plane, the corresponding projection is marked as 'o'. Escape points with projections to the left of the separating curve tend to start converging trajectories, while those projected to the right of the curve initiate an excursion towards Belief 3. There are several points well to the left of the separating curve which nevertheless start excursions. We believe that these points represent very unlikely escapes, during which the structure of the matrix R changes a lot, and these points are actually very far from $\widetilde{v}_1$ in the original 27D space.

Using a thus constructed boundary improves the match to simulations for gain value $\epsilon = 0.001$. The theoretical escape calculated in this way occurs through the point that lies in the approximate direction (1,-4)´ in $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ coordinates. The distance between the average of empirical escape points and the theoretical escape point equals $d2 = 31\%$ of the distance between the SCE and the theoretical point. Figure 2.6 shows that there is some accumulation of the escape tangents towards -4; however, the majority tends to center on -5, exactly as in the case of escape from the cylinder.

According to point (b) of the Theorem A.1, as $\epsilon \to 0$, the probability of observing the escape in the neighborhood of the point minimizing action functional converges to one. This is not the behavior observed in our simulations: For both boundaries used above simulated escapes tend to occur close to points in (1,-5)´ direction rather than the theoretically predicted (1,-6.5)´ or (1,-4)´.

The source of the discrepancy lies in the non–validity of continuous–time approximation for the values of ϵ used in economic literature and for simulations in the comparison above. To see it, one could compare the mean dynamics given by (2.11) and the realizations of the stochastic process simulated in (2.6). As is mentioned above, for θ near the SCE such that the mean dynamics points towards $\overline{\theta}$, mean dynamics magnitude $(R^{-1}E[g(\gamma, \xi_n)])$ is dramatically lower than the typical realizations of $R^{-1}g(\gamma, \xi_n)$: the "noise–to–signal" ratio, $\frac{\left\|R^{-1}(g(\gamma,\xi_n)-E[g(\gamma,\xi_n)])\right\|}{\|R^{-1}E[g(\gamma,\xi_n)]\|}$, ranges from hundreds to several thou-

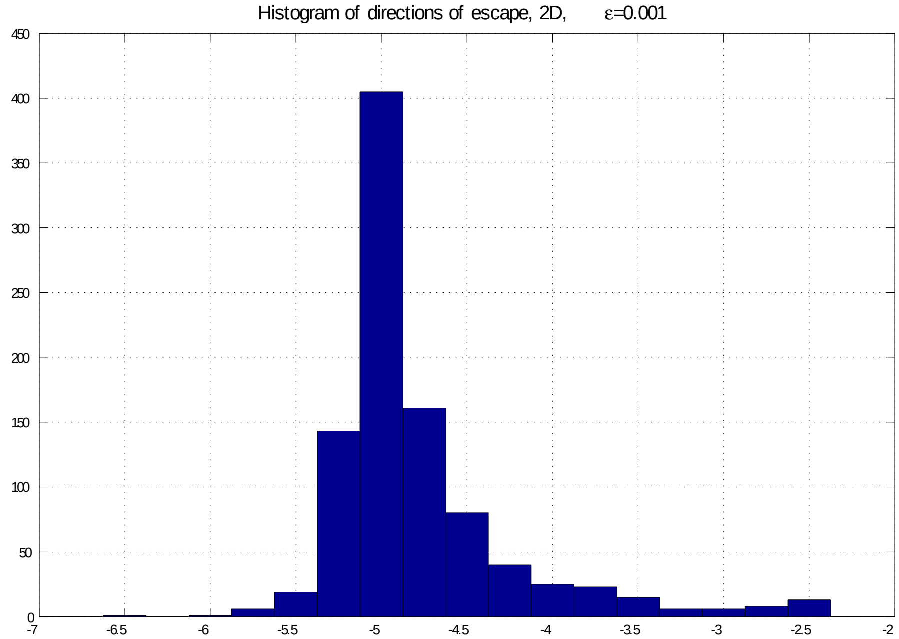

Figure 2.6: Histogram of tangents of escape directions in (γ_1, γ_2) space for $\epsilon = 1 \cdot 10^{-3}$. Escapes from the region bounded by the "separating curve."

sand, depending on θ. Figure 2.3 plots this ratio for points along the eigenvector $\widetilde{v}_1$, averaged over 4000 realizations of (W_{1n}, W_{2n}) for every point. The horizontal axis shows corresponding $\widetilde{\gamma}_1$ values. Near the SCE, the noise–to–signal ratio tends to be extremely high. In this situation one expects (2.16) to be a good approximation of (2.6) only if the system stays in the neighborhood of every point θ long enough to allow the average of $R^{-1}g(\gamma, \xi_n)$ to approach $R^{-1}E[g(\gamma, \xi_n)]$. With noise–to–signal ratios from 10 to 1000, this means hundreds and thousands of iterations near **every** point θ. However, for the values of ϵ used in CWS and commonly applied in the literature (ϵ=0.001÷0.01), the expected escape time is measured in hundreds of iterations: this is the time spent by the system near **all** points around the SCE. Therefore, ϵ=0.001÷0.01 is not small enough for the approximation (2.16) to be valid.

Given such large noise–to–signal ratio, one could simply disregard the mean dynamics (set $D_\theta p(\theta) = 0$) and repeat minimization of the action functional. This is our third way of deriving escape dynamics. The region D is the cylinder described above. The theoretical results of the problem of minimizing action functional on the cylinder

ϵ	Dynamic model d1, %	d2, %	d3, %	Static model d1, %	d3, %
$2 \cdot 10^{-5}$	75.45	86.72	99.99		
$3 \cdot 10^{-5}$	81.56	73.65	72.25	24.82	244.93
$5 \cdot 10^{-5}$	88.50	70.69	40.95	29.37	229.23
$1 \cdot 10^{-4}$	93.04	57.73	20.39	40.91	189.36
$2 \cdot 10^{-4}$	92.78	45.68	22.75	54.73	141.64
$4 \cdot 10^{-4}$	95.83	35.26	7.61	67.41	97.85
$1 \cdot 10^{-3}$	99.65	30.98	4.35	81.02	50.82
$1 \cdot 10^{-2}$	99.39	127.86	11.87	93.78	6.77

Table 2.1: Distance between the average of simulated escape points and the theoretically predicted escape points, expressed as a percentage of the distance between the SCE and the theoretically predicted point.

in the case of diffusion without drift term are presented in Appendix A.2. Theoretical escape occurs in the direction of the largest eigenvalue of Σ, $\tilde{v}_1$.[18] This way provides a much better agreement between the theory and the simulations: for the same radius of the cylinder as in the first approach, distance $d3$ between the mean of observed escape points and the theoretically predicted point is only 4.4% of the distance between the SCE and the theoretical point. Figure 2.4 shows that most runs end in escapes along the direction (1,-5)´, which is the theoretically predicted one for this approach.

To support further our claim that $\epsilon = 0.001$ is not low enough to guarantee sufficient averaging, we have performed simulation runs for smaller values of ϵ. Consider Figure 2.7, which plots a histogram of escape directions from the cylinder, for a 1000 simulations with $\epsilon = 2 \cdot 10^{-5}$. Comparison with Figure 2.4 shows that one indeed observes an accumulation of escape directions towards the theoretically predicted direction of -6.5, but there is still a long way to go: A full 14% of escapes occur in the bin centered on -5 with the width of 0.25, and only 66% escape direction tangents are in bins with centers below -5. Given that theoretically one expects a mode at -6.5, we calculated the share of escapes in the bin centered at -6 and below it. Only 9.6% of escapes fall into this category, which is still much better than less than 1% observed for $\epsilon = 1 \cdot 10^{-3}$.

In contrast, performance of the second way of selecting ∂D does not improve as ϵ decreases.[19] At $\epsilon = 1 \cdot 10^{-3}$, only 41% of escapes are at -5, and 26% escape at "correct"

[18] This escape direction could be derived more easily: take the largest eigenvector of R^{-1} and project it into $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space. The result is the same as the result based on the formula in Appendix A.2 up to third decimal point.

[19] The reason for the failure of the second way for smaller ϵ is clear: for large ϵ, the majority of escapes are

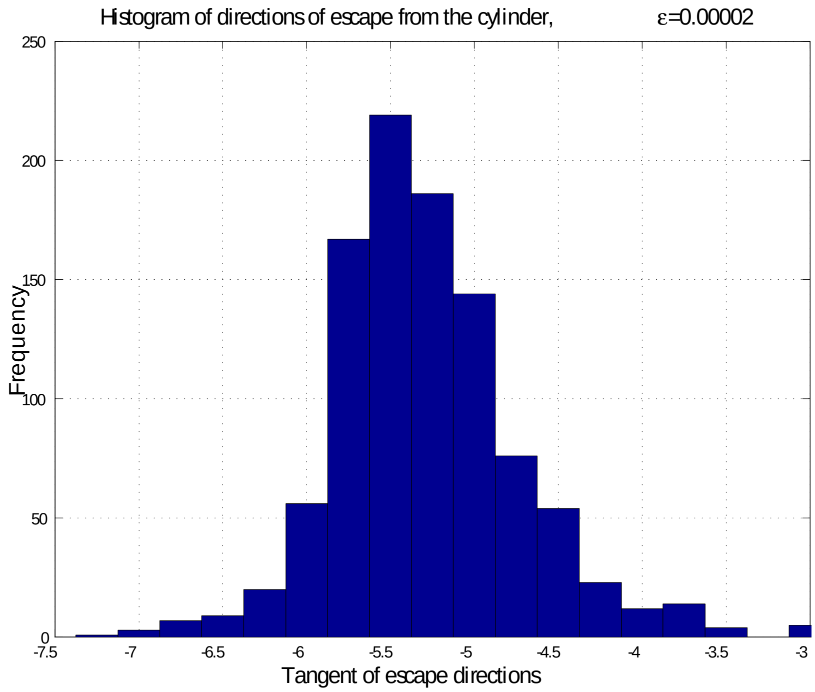

Figure 2.7: Histogram of tangents of escape directions in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space for $\epsilon = 2 \cdot 10^{-5}$. Cylinder with radius described in the text.

-4.5 or above (recall that the second way predicts an escape direction of -4). However, the percentage of "correct" escapes does not increase as ϵ falls but fluctuates between 25% and 28% for $\epsilon = 2 \cdot 10^{-4}$, $4 \cdot 10^{-4}$, and $2 \cdot 10^{-5}$. A large number of escapes occurs below -5 and below -6 (59% and 9.7% respectively at $\epsilon = 2 \cdot 10^{-5}$). In other words, the distribution of the directions of escape from the region bounded by the separating curve resembles the distribution of escapes from the cylinder, even though theoretically we expected them to diverge.

concentrated along $\widetilde{v}_1$, and we used a point on $\widetilde{v}_1$ as a starting point in deriving the trajectory which, after being projected into $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ plane, became the separating curve. As a result, we are relatively confident that this curve is a good description of the true 27–dimensional separating surface for majority of escaping trajectories. As ϵ decreases, more and more escapes start to take place away from $\widetilde{v}_1$ (and thus away from the initial point used to derive the separating curve), and the curve stops working as a result. We could use a point on the vector $T_1\overline{G}^{1/2}\xi$ derived in Appendix A.2 as the initial point for deriving the separating curve. We believe that this trajectory, projected into $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ plane, would have worked well for very small ϵ, when almost all escapes do occur in the direction prescribed by $\overline{G}$.

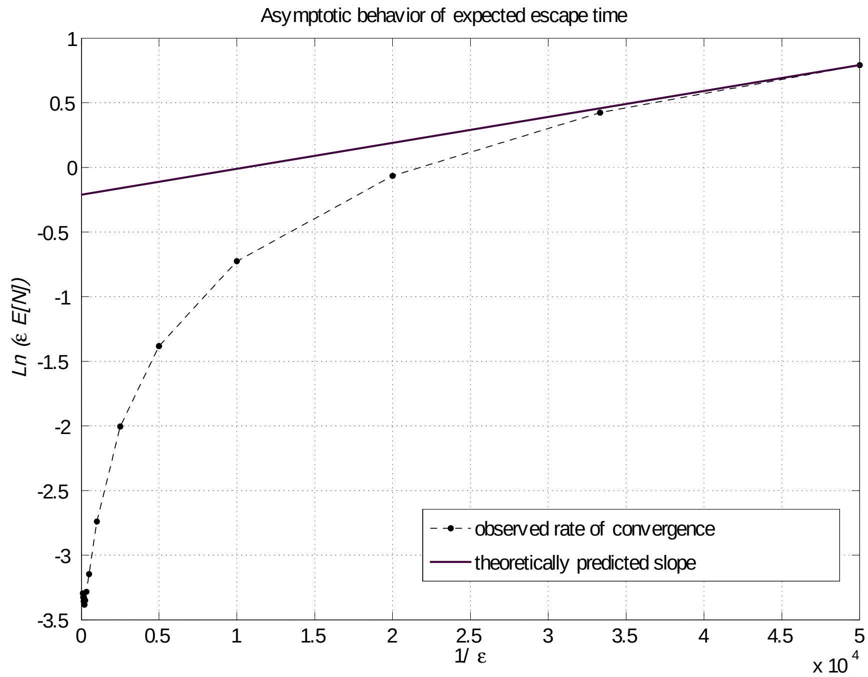

Figure 2.8: Comparison of empirical behavior of average escape time (dots) and the theoretical prediction (solid line).

2.4.3 Analytical Results vs. Simulations: Escape Time

To see how the proposed approach works in terms of predicting the mean escape time we exploited the relation (2.17). It implies that a plot of $\ln E_x\left(\tau^{\varepsilon}\right)$ vs. $1/\epsilon$ is a straight line with the slope equal to the rate of convergence $\overline{I}$. Note that τ^{ε}, first escape time, is given in continuous time units of the approximating diffusion (2.16), and is approximately equal to ϵ times the expected number of iterations of the discrete–time process (2.10) needed to observe the first escape. Figure 2.8 shows that the straight line is not observed: The slope decreases as one moves towards higher $1/\epsilon$ (lower ϵ), and seems to converge asymptotically to the solid line with the slope $\overline{I}$ only for the lowest considered values of ϵ .[20]

Table 2.2 shows the theoretically predicted values of the slope $\overline{I}$ for the first and second way of selecting ∂D as well as empirically observed slopes (calculated at the slope of

[20]Though the limiting characteristics of mean escape time predicted by the large deviations theory do not hold true for economically plausible region of gain ϵ, the empirical distribution of escape times follows the theoretically predicted exponential distribution even for “large” ϵ: Figure 2.9 shows that the logarithm of the empirical cumulative distribution function, $ln[\Pr(\tau^{\varepsilon} \geq T)]$, is approximately linear in T, as expected for the exponential distribution.

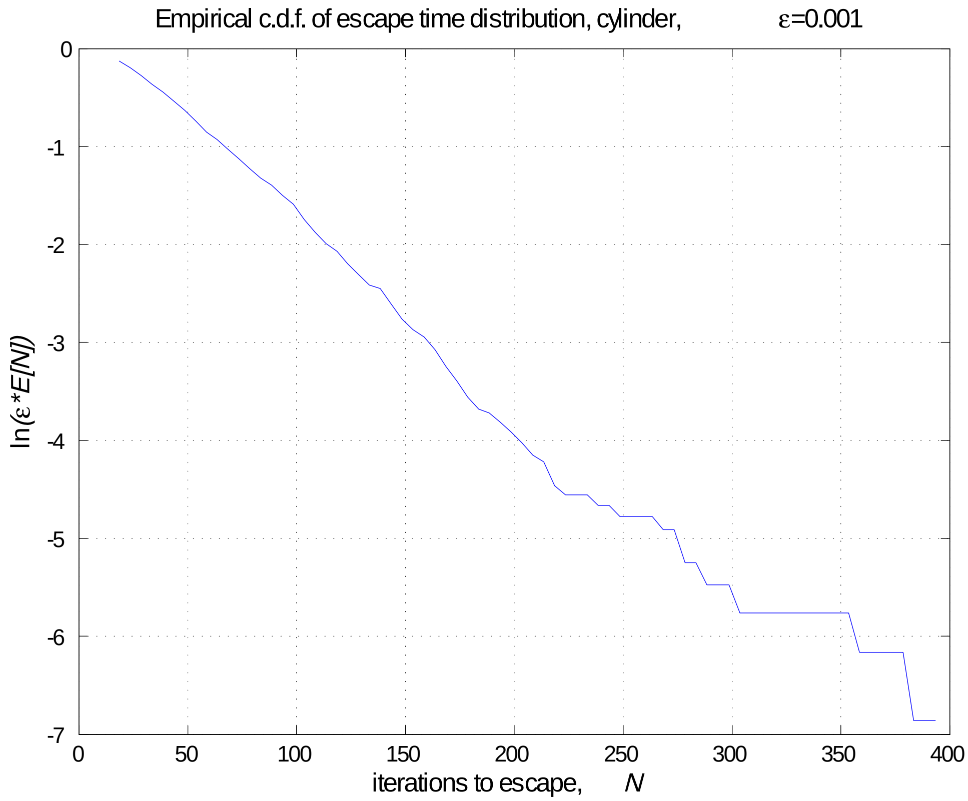

Figure 2.9: Cumulative distribution function of escape times

the line connecting the last two observations in Figure 2.8 and its analog for the separating curve). For the first (second) way, the empirical slope almost converges to (undershoots) the theoretically predicted value for $\epsilon = 2 \cdot 10^{-5}$. As explained in footnote 26, our method is likely to produce values of $\bar{I}$ which are *higher* than the true ones, which explains the observed undershooting.

The theoretical formula for mean escape time in the third way of characterizing escape dynamics in the model can be derived using the formula for mean exit time of one–dimensional Brownian motion in Karatzas and Shreve [46]. We derive the mean escape time for the projection of the process $d\varphi_t = \sqrt{\epsilon}\Sigma^{1/2}(\bar{\theta})dW_t$ on the most probable direction of escape, the "largest" eigenvector of Σ. The formula for the mean escape time, derived in Appendix A.3, is given as $E\tau^{\varepsilon} = \frac{rad^2}{\epsilon\lambda}$, where λ is the largest eigenvalue of Σ, and rad is the distance between the SCE and the point where the "largest" eigenvector of Σ crosses the cylinder described above. In Table 2.3, we compare this formula's predictions with averages from simulations. The formula performs very well, especially for $\epsilon \in [0.001; 0.01]$,

	ϵ	Dynamic model Way 1	Dynamic model Way 2	Static model Way 1
Simulations	$2 \cdot 10^{-5}$	$2.21 \cdot 10^{-5}$	$2.14 \cdot 10^{-5}$	
	$3 \cdot 10^{-5}$	$3.67 \cdot 10^{-5}$	$3.19 \cdot 10^{-5}$	$1.96 \cdot 10^{-4}$
	$5 \cdot 10^{-5}$	$6.60 \cdot 10^{-5}$	$4.92 \cdot 10^{-5}$	$1.86 \cdot 10^{-4}$
	$1 \cdot 10^{-4}$	$1.31 \cdot 10^{-4}$	$8.46 \cdot 10^{-5}$	$2.30 \cdot 10^{-4}$
	$2 \cdot 10^{-4}$	$2.49 \cdot 10^{-4}$	$1.77 \cdot 10^{-4}$	$3.42 \cdot 10^{-4}$
	$4 \cdot 10^{-4}$	$4.90 \cdot 10^{-4}$	$3.79 \cdot 10^{-4}$	$5.66 \cdot 10^{-4}$
	$1 \cdot 10^{-3}$	$8.14 \cdot 10^{-4}$	$6.76 \cdot 10^{-4}$	$1.32 \cdot 10^{-3}$
Theory		$2.01 \cdot 10^{-5}$	$3.01 \cdot 10^{-5}$	$3.20 \cdot 10^{-4}$

Table 2.2: A comparison of the theoretically derived value of the action functional and empirically observed slope of $\ln E_x(\tau^\varepsilon)$ vs. $1/\epsilon$ line.

the range used previously by CWS and others. Indeed, when we plot average simulated escape time vs. $1/\epsilon^2$ in Figure 2.10,[21] we see a straight line with the slope approximately equal to $\frac{rad^2}{\lambda}$ for a large range of ϵ. The formula starts to lose precision once one moves to lower ϵ, *i.e.* into the region where the averaging is better as the system spends more periods in the neighborhood of SCE, and so the large deviations theory becomes to be more applicable to characterizing mean escape times.

.

Simulation results for very low ϵ tell a consistent story: one can use large deviations theory estimates for escape time only for gain values $\epsilon \lesssim 2 \cdot 10^{-5}$; even $\epsilon = 2 \cdot 10^{-5}$ is still not low enough to observe the limiting behavior predicted by the large deviations theory. Note that this is true for both continuous–time approach developed here and discrete–time approach used by CWS.[22] As for the continuous–time approach, one can avoid the non–applicability of the large deviations theory for economically interesting values of ϵ by disregarding the mean dynamics altogether and following the third way of deriving the escape dynamics, which works well for such ϵ.

For the lowest ϵ considered here, the large deviations theory just begins to work. Is it possible to claim that empirically observed disinflations are, indeed, escapes from

[21] Observe that τ^ε, first escape time, is given in continuous time units of the approximating diffusion (2.16), and is approximately equal to ϵ times the expected number of iterations of the discrete time process (2.10) needed to observe the first escape (see the beginning of this subsection for explanation). Therefore, we divide the "continuous–time mean escape time" $E\tau^\epsilon$ by ϵ to get the "mean number of periods before escape."

[22] Even if one believes that our calculations overestimate the rate of convergence because of the arguments presented in footnote 25 and the true $\overline{I}$ is closer to the numbers given in CWS, the empirically observed slope at $\epsilon = 0.001$ is 80 times larger than predicted. This result means that both discrete– and continuous–time approaches fail at least for $\epsilon \sim 0.001$.

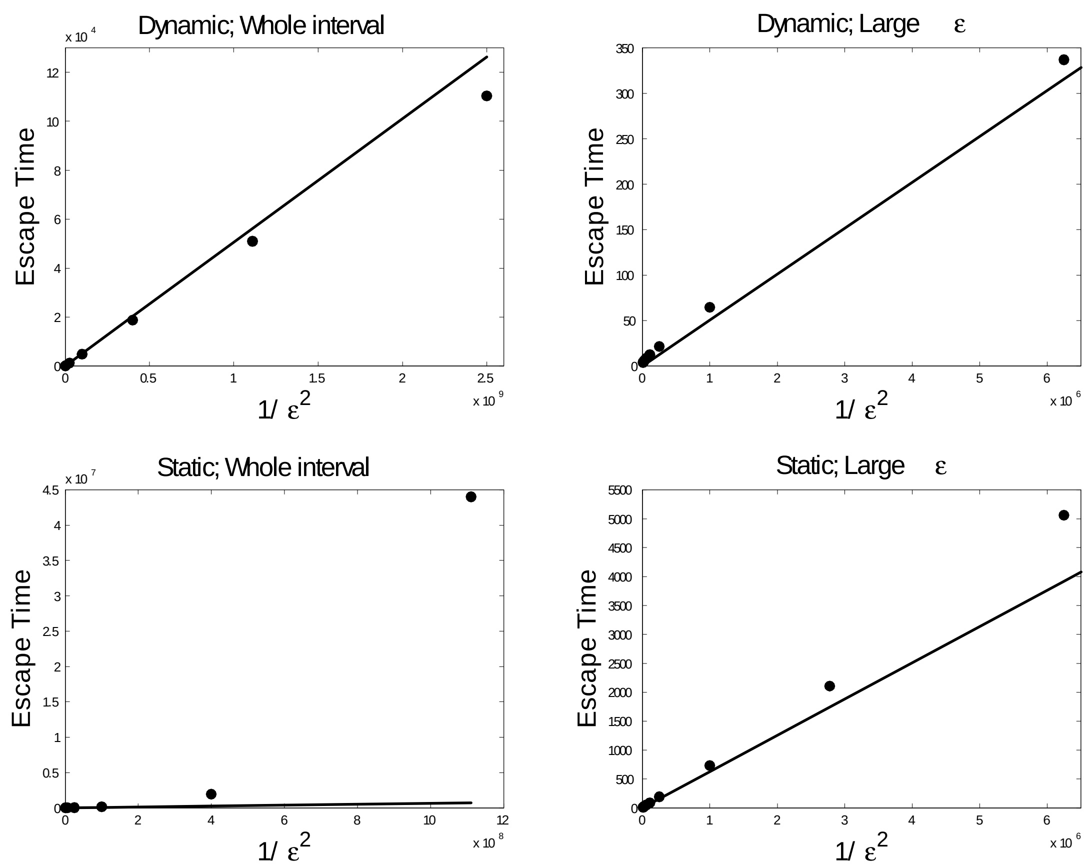

Figure 2.10: Expected escape times predicted for a one–dimensional Brownian model without drift versus simulation results.

the SCE generated by the model? At $\epsilon = 2 \cdot 10^{-5}$, the average number of simulation periods needed to observe the first escape is about $1 \cdot 10^5$ (for cylinder) and $1.5 \cdot 10^5$ (for the separating curve). Add to these numbers the time needed to travel to the induction hypothesis plane $\widetilde{\gamma}_1 = 0$ (of order 10^4) and recall that the time period in the Phelps model could not be much lower than a quarter. In this economy, one would wait, on average, twenty thousand years or longer for the low inflation episode caused by adaptive learning with constant gain equal to $2 \cdot 10^{-5}$. It is immediately obvious that the region of ϵ values, for which large deviations theory estimates of mean escape time start to be applicable in the dynamic model of CWS, is far removed from those ϵ values which lead to simulated escapes at empirically interesting times, such as 65 periods at $\epsilon = 1 \cdot 10^{-3}$.

	Dynamic model		Static model	
ϵ	Simulations	Theory, $\frac{rad^2}{\epsilon^2\lambda}$	Simulations	Theory, $\frac{rad^2}{\epsilon^2\lambda}$
$2 \cdot 10^{-5}$	$1.10 \cdot 10^5$	$1.26 \cdot 10^5$		
$3 \cdot 10^{-5}$	$5.10 \cdot 10^4$	$5.61 \cdot 10^4$	$4.38 \cdot 10^7$	$6.97 \cdot 10^5$
$5 \cdot 10^{-5}$	$1.88 \cdot 10^4$	$2.02 \cdot 10^4$	$1.98 \cdot 10^6$	$2.51 \cdot 10^5$
$2 \cdot 10^{-4}$	$1.26 \cdot 10^3$	$1.26 \cdot 10^3$	$1.56 \cdot 10^5$	$6.27 \cdot 10^4$
$4 \cdot 10^{-4}$	336.96	315.65	4928.00	3919.27
$1 \cdot 10^{-3}$	64.59	50.50	733.57	627.08
$2 \cdot 10^{-3}$	21.49	12.63	189.98	156.77
$3 \cdot 10^{-3}$	12.50	5.61	87.00	69.68
$4 \cdot 10^{-3}$	8.77	3.16	52.08	39.19
$5 \cdot 10^{-3}$	6.79	2.02	34.39	25.08
$6 \cdot 10^{-3}$	5.99	1.40	24.76	17.42
$7 \cdot 10^{-3}$	4.98	1.03	19.14	12.80
$8 \cdot 10^{-3}$	4.49	0.79	15.02	9.80
$9 \cdot 10^{-3}$	4.12	0.62	13.32	7.74
$1 \cdot 10^{-2}$	3.70	0.51	11.16	6.27

Table 2.3: A comparison of the theoretically derived values of expected escape time and empirically observed average escape times.

2.5 Discussion

2.5.1 Better Averaging for Larger ϵ

As we have shown in the dynamic model of CWS at empirically relevant values of the constant gain parameter, the mean escape time can be easily characterized using a simple formula for the expected exit time of one–dimensional Brownian motion, while the large deviations theory predictions of mean escape time do not hold. This is due to three facts. A large value of u leads to very large λ_1/λ_2, the ratio of the two largest eigenvalues of the inverse second moments matrix at SCE. This fact makes the SRA dynamics almost one–dimensional. Second, both static and dynamic models of CWS have a very specific phase portrait: despite potentially global stability of the SCE, the region where the mean dynamics points back to it is exceedingly small. Third, the dominant eigenvalue λ_1 is huge, which means that any stochastic deviation in the "only" direction is strongly magnified. A combination of these three features of the SRA in the dynamic CWS model makes "escape" easy and puts very tight requirements on the values of ϵ used in the constant gain learning algorithm. In particular, this means that values of ϵ commonly used in the literature are not small enough to guarantee enough time for averaging, and therefore, methods of characterizing the mean escape time, based on the theory of large deviations which relies on the mean (averaged) dynamics, are not expected to work well in this particular setting. This argument is applicable to both the continuous–time approach and to the discrete–time approach used in CWS and elsewhere. Nevertheless, a version of the continuous–time approach which disregards the mean dynamics, and in fact does not rely on the large deviations theory to characterize the mean escape time, provides a very good fit to the simulations.

Lack of time for averaging near the SCE leads to the relative failure of our first and second ways in deriving the properties of escapes. This failure could be tracked down to the limited applicability of (2.11) and of approximating diffusion (2.16). With better averaging, both large deviations theory characteristics of mean escape time and continuous–time approximation (2.16) improve. We have three suggestions on how to achieve better averaging in this model for empirically interesting region of ϵ.

First, the matrix $\overline{R}^{-1}$ should be more balanced: A lower value of λ_1/λ_2 effectively increases the dimensionality of the problem and expands the volume of the state space available to the system, thus increasing the expected escape time and producing better averaging. In the current model, a more balanced second moments matrix means lower value of u. Second, having a stronger drift towards SCE under mean dynamics (this probably implies a larger region of immediate attraction to the SCE) will help to achieve better averaging around SCE and thus ensure that (2.16) approximates (2.6) reasonably well. We conjecture that smallness of the region where the mean dynamics points towards SCE is due to the fact that at the SCE, the learning is not well specified: some of the regressors are perfectly collinear in both static and dynamic models of CWS. In other words, forcing the agents to use better specified learning might help. Third, reducing the magnitude of λ_1 while keeping fixed the SCE's region of attraction increases the time spent by the system in this region and provides better averaging, at a potential cost of a higher expected escape time; as a result, the region of the constant gain parameter ϵ where large deviations theory's asymptotic predictions start to be valid is more likely to not include empirically interesting magnitudes of ϵ.

2.5.2 Static Model vs. Dynamic Model

In order to test some of these conjectures, we repeated our analysis for the static model of CWS. In contrast to the dynamic model analyzed previously, the government's beliefs do not take into account lagged inflation and unemployment rates, and the vector X_{n-1} contains only the constant.[23] We expect the large deviations theory to become applicable for larger values of ϵ than with the dynamic model because the phase portraits of the mean dynamics in the dynamic (in $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$ coordinates) and static models are very similar, but the largest "static" eigenvalue of $\overline{R}^{-1}$ equals 26.09 and is much smaller than 3084 for the dynamic model. On the other hand, the ratio of the first two eigenvalues, λ_1/λ_2, equals 7561 compared to 106. In other words, the static model is even more one–dimensional than the dynamic one, but the system is expected to spend more time in the neighborhood of the SCE before the first escape and thus achieve better averaging than in the dynamic model.[24]

Based on the discussion in the previous section, we do not consider the second way of selecting ∂D and use only the first and the third one. The cylinder is chosen so that it intersects the "largest" eigenvector $\widetilde{v}_1$ at γ_1 =-0.975, using the same reasoning as in the dynamic model case. Theoretically predicted directions of escape out of the cylinder are (1,-7.5)´ and (1,-5)´ for the first and third way, respectively.

On several dimensions, the first way produces more favorable results when applied to the static model. First, at $\epsilon = 1 \cdot 10^{-3}$, only 22% of escapes happen in the -5 bin, compared to 56% for the dynamic model. At $\epsilon = 2 \cdot 10^{-4}$, this number drops to 6.7% for the static model, but still stands at 49% in the dynamic one. At $\epsilon = 3 \cdot 10^{-5}$ in the static model, a full 34% of the escapes happen at the direction -7 or below and 100% of them are below -5. As is clear from the Table 2.1, the distance to the escape point predicted by the first way declines very fast as ϵ decreases in the static model, but it decreases only slightly in the dynamic one. Finally, already at $\epsilon = 1 \cdot 10^{-4}$, the empirically observed slope

[23] The government's problem (2.3) can be solved explicitly for the policy function $h(\gamma)$. A preliminary analysis of the static CWS model was performed in Evans and Honkapohja [29, Section 14.4]. In particular, the matrix V derived on p. 358 is nothing else but the Gramian $\overline{G}$ of the 2–dimensional problem which discards 3 elements of the covariance matrix. As stated by Evans and Honkapohja [29], one could do so because the matrix $D_\theta p(\bar{\theta})$ is block–lower–triangular and the matrix $\Sigma(\bar{\theta})$ is block–diagonal. In order to preserve continuity with our analysis of the dynamic model, we analyze the full 5–dimensional static problem, with two elements of the vector γ and 3 elements of the variance–covariance matrix R.

[24] This prediction is borne out by the simulations: at $\epsilon = 1 \cdot 10^{-2}$, $1 \cdot 10^{-3}$, and $2 \cdot 10^{-4}$, the average escape times out of the cylinder which intersects $\widetilde{v}_1$ at -0.985 are 3.7, 64.6, and 1255 periods for the dynamic model, while for the static one the corresponding numbers are 5.6, 265.6, and 6903 periods, respectively.

of the $\ln E_x(\tau^\varepsilon)$ vs. $1/\epsilon$ line is lower than the theoretically predicted one, see Table 2.2.[25]

As for the third way we see that it works well in predicting the point of escape (see Table 2.1) and in predicting the mean escape time (see Table 2.3). However, the simplified continuous–time approximation (without drift term) and the simple mean escape time formula based on it are valid for the region of ϵ shifted to a larger ϵ compared to the one for the dynamic model, see Figure 2.10. This is explained by better averaging for larger ϵ in the static version of the model.

2.5.3 Comparisons with CWS

Comparing the predictions of our most successful (third) way of deriving the most probable escape point and the mean escape time *for economically plausible region of gain values* in both dynamic and static versions of the CWS model with the CWS predictions, we obtain that our estimates of *magnitudes* of mean escape times fit simulations very well, whereas the mean escape times' *limiting behavior* predicted by the large deviations theory used by CWS is not confirmed by simulations. The reason for this failure (and for the failure of our first and second ways of describing mean escape time for economically interesting gain values) is bad averaging as discussed above.

In terms of the escape point, all three ways used in this paper have the same escape point as understood by CWS[26]: escaping trajectories hit the "induction hypothesis" plane very close to the point prescribed by the largest eigenvalue's eigenvector of $\overline{R}^{-1}$, the inverse second moments matrix of beliefs evaluated at the SCE. We explain this by the fact that this "escape" is for the most part a *deterministic* movement along the mean dynamics trajectory, and all such trajectories are very close to each other, see Figure 2.2.

We have to address three technical considerations that could potentially influence our comparison. The first is "Kushner critique": Can one guarantee that the escape dynamics generated by a continuous–time approximating process is a valid approximation of the escape dynamics of the original discrete–time learning process? Kushner [52] pro-

[25] Recall that $\overline{I}$, which determines this slope, is called quasipotential. It measures the energy needed to get the system out of region D around the point O. The reason one needs to spend the energy at all is because of the drift, or non–stochastic component of the diffusion, which points back towards O. We have selected our region D in such a way that the drift pointing inwards becomes very weak near the boundary ∂D. However, as the approximation to (2.6) is made at point O, it overestimates the strength of the drift which has to be overcome in order to cross ∂D, and so overestimates $\overline{I}$. The argument from Dembo and Zeitouni [22], cited in Section 2.1, persuades us that this upward bias is not likely to be very large.

[26] In their paper they consider escape out of a cylinder with radius 66 times as large as in our paper.

posed that this question could be answered by checking whether the action functional of a discrete–time process converges to the action functional of its continuous–time approximation. Checking the convergence is very hard to do taking into account that the action functional for a discrete–time process in the CWS model depends on a numerically derived function. Therefore, the only way of judging the performance of the large deviations theory approach based on a continuous–time approximation of the discrete–time SRA is to compare the predictions of this approach to the predictions of the discrete–time approach and to simulation results.

Second, the CWS approach ignores the cross effects of the second moments matrix R for dynamic version of the model due to numerical complications, while we fully take them into account. We estimate that the influence of this assumption is not large as the matrix Σ is close to being block diagonal (see footnote 13 for a discussion). We believe that if one were to implement the CWS approach without ignoring the cross effects of R, the results obtained would be more valid than those obtained here. This path, however, might be blocked by computational complications.

Third, the discrete–time version of large deviations theory does not contain theoretical results for the most probable point of escape and mean escape time in case of unbounded (for example, Gaussian) shocks (see CWS, Theorem 5.3). The result which is available for unbounded shocks is that "the probability of observing an escape episode is exponentially decreasing in the gain with the rate given by the minimized value of the cost function $\overline{S}$", see CWS, p. 13. The extent to which this lack of theoretical results influences numerical predictions is not generally known, but has been shown to be small in some situations: CWS use both normal (unbounded) and binomial (bounded) shocks in the static model and obtain the values of $\overline{S}$ ($\overline{I}$ in our notation) which are numerically very close to each other; predicted most probable escape paths also are very similar.

2.6 Conclusion

We extended a continuous–time approach for the analysis of escape dynamics in economic models with adaptive constant gain learning. Foundations of this approach were laid down by Evans and Honkapohja [29, Ch. 14], Williams [64], and Kasa [47]. This approach is based on applying the results of FW's continuous–time version of large deviations theory to the diffusion approximation of the original discrete–time learning process.

When applied to the Phelps problem of government controlling inflation using an approximate Phillips curve, deriving escape dynamics characteristics for the "original" diffusion approximation with a drift term did not generate results compatible with the results of simulations in terms of the mean escape time. This is due to the limited validity of such an approximation for economically plausible values of the constant gain parameter ϵ. Limited validity of the approximation, in turn, is caused by bad averaging in the CWS model. To account for high "noise–to–signal" ratio near the SCE, we used a "modified" diffusion approximation without the drift term and the formula for the mean exit time of one–dimensional Brownian motion, rather than limiting characteristics provided by the theory of large deviations. We managed to predict the values of mean escape times with high precision for the "modified" approximation.

All our ways of deriving escape dynamics characteristics work well in predicting the final point of escape: Escape occurs in the small neighborhood of the SCE, then mean dynamics move the system along the largest eigenvector of the inverse second moments matrix evaluated at the SCE towards the "induction hypothesis" plane. As for predicting the most probable point of "initial" escape out of a small neighborhood of the SCE, the ways based on an "original" continuous–time approximation did not work well for economically plausible region of ϵ because the approximation has limited validity at such ϵ. Considering an escape out of a separating surface on which the mean dynamics changes its direction, thus taking into account some information about the behavior of the mean dynamics away from the SCE, provides better results for the point of "initial" escape.

As another result of this paper we express reservations regarding the applicability of large deviations theory for the characterization of mean escape time for economically plausible values of gain in both versions of the CWS model. We show that for the region of gain values used in economics literature, simple considerations and formulae work much better than large deviations theory's results. This, again, is explained by bad averaging for a relatively large ϵ in this model.

We suggest two changes to help the approaches based on large deviations theory work better in terms of characterizing mean escape time for the model and gain values considered: to set lower mean unemployment rate, in order to construct a more balanced second moments matrix, and to use better specified learning of agents. The same changes will help our "original" continuous–time approximation to become valid for larger gain values. In general, one has to look for economically sensible models with better averaging for economically plausible gain values in order to apply large deviations theory characteristics of the mean escape time.

Finally, we believe that utilizing a continuous–time approximation can be used to analyze escape dynamics in more complicated models, where it is not possible to derive analytical characteristics of escape dynamics in discrete time. For example, the model with a dynamic Phillips curve can be studied as a possible extension of the approach proposed here. The question, however, remains whether large deviations theory predictions of mean escape time would be valid in this model for economically plausible ϵ, or whether one would have to employ something resembling our reliance on the "modified" approximation and the mean exit time result for one–dimensional Brownian motion. This is the focus of our current research.

Chapter 3

Stochastic Gradient versus Recursive Least Squares Learning

Stochastic Gradient versus Recursive Least Squares Learning

Sergey Slobodyan, Anna Bogomolova, and Dmitri Kolyuzhnov*

CERGE–EI†
Politických vězňů 7, 111 21 Praha 1,
Czech Republic

Abstract

In this paper, we perform an in–depth investigation of the relative merits of two adaptive learning algorithms with constant gain, Recursive Least Squares (RLS) and Stochastic Gradient (SG), using the Phelps model of monetary policy as a testing ground. The behavior of the two learning algorithms is very different. Under the mean (averaged) RLS dynamics, the Self–Confirming Equilibrium (SCE) is stable for initial conditions in a very small region around the SCE. Large distance movements of perceived model parameters from their SCE values, or "escapes," are observed.

On the other hand, the SCE is stable under the SG mean dynamics in a large region. However, actual behavior of the SG learning algorithm is divergent for a wide range of constant gain parameters, including those that could be justified as economically meaningful. We explain the discrepancy by looking into the structure of eigenvalues and eigenvectors of the mean dynamics map under SG learning.

Results of our paper hint that caution is needed when constant gain learning algorithms are used. If the mean dynamics map is stable but not contracting in every direction, and most eigenvalues of the map are close to the unit circle, the constant gain learning algorithm might diverge.

JEL Classification: C62, C65, D83, E10, E17
Keywords: constant gain adaptive learning, E–stability, recursive least squares, stochastic gradient learning

* *{Sergey.Slobodyan, Anna.Bogomolova, Dmitri.Kolyuzhnov}@cerge-ei.cz.*

† CERGE–EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic.

3.1 Introduction

In this paper, we perform an in–depth investigation of the relative merits of two adaptive learning algorithms with constant gain, Recursive Least Squares (RLS) and Stochastic Gradient (SG). Properties of RLS as a learning algorithm are reasonably well understood as it has been used extensively in the adaptive learning literature. For an extensive review, see Evans and Honkapohja [29]. SG learning received more limited attention in the past, but the situation is changing: Evans, Honkapohja and Williams [32] promote the constant gain SG (and generalized SG) as a robust learning rule, which is well suited to the situation of time–varying parameters.

A different motivation for studying the properties of the SG learning comes from recent interest in heterogeneous learning (cf. Honkapohja and Mitra [42] or Giannitsarou [37]). In this literature, several types of agents use different adaptive learning rules to arrive at the parameter values of the model. Often, some of the groups are using RLS while the others employ SG. A desirable property of such a model is its stability under all implemented types of learning.

Finally, our interest is not restricted to the dynamics of the learning algorithm in a small neighborhood of the rational expectations equilibrium (REE) which motivates our focus on constant gain learning. It is known that E–stability of the REE, which implies local stability under RLS learning with decreasing gain, does not automatically imply local stability under SG with decreasing gain, see Giannitsarou ([38]). Here the equilibrium is E–stable under both RLS and SG learning, but the behavior of the constant gain versions of the two methods is substantially different away from the equilibrium.

As a testing ground for comparison, we use the Phelps problem of a government controlling inflation while adaptively learning the approximate Phillips curve, studied previously by Sargent [61] and Cho, Williams, and Sargent [17] (CWS hereafter). A phenomenon known as "escape dynamics" can be observed in the model under the constant gain RLS learning. In Kolyuzhnov, Bogomolova, and Slobodyan [50], we applied a continuous–time version of the large deviations theory to study the escape dynamics and argued that a simple approximation by a one–dimensional Brownian motion can be better suited for describing the escape dynamics in a large interval of values of the constant gain. Here, we derive an even better one–dimensional approximation and discuss the Lyapunov function–based approach in establishing the limits of applicability of this approximation.

We also extend our analysis to the SG constant gain learning.

The rest of the paper is organized as follows. We briefly describe the dynamic and static versions of the model of CWS and define the RLS and SG learning in Section 2. In Section 3, we present and contrast the non–local effects arising under the constant gain versions of these algorithms and discuss the possible explanations for the difference in behavior of the mean dynamics and the actual real–time learning algorithm. Section 4 concludes.

3.2 The Model and Learning Algorithms

The economy consists of the government and the private sector. The government attempts to minimize losses from inflation π_n and unemployment U_n:

$$\min_{\{x_n\}_{n=0}^{\infty}} E\sum_{n=0}^{\infty}\beta^n\left(U_n^2+\pi_n^2\right), \tag{3.1}$$

It uses the monetary policy instrument x_n to control π_n, Eq. (3.2b). It believes (in general, incorrectly) in the Phillips curve (3.2c). The true Phillips curve is given by (3.2a): Unemployment is affected only by unexpected inflation. The private sector possesses rational expectations $\widehat{x}_n = x_n$ about the inflation rate, and thus unexpected inflation shocks come only from monetary policy errors. The whole model is presented below.

$$U_n = u-\chi\left(\pi_n-\widehat{x}_n\right)+\sigma_1 W_{1n},\ u>0,\ \theta>0, \tag{3.2a}$$

$$\pi_n = x_n+\sigma_2 W_{2n}, \tag{3.2b}$$

$$U_n = \gamma_1\pi_n+\gamma_{-1}^T X_{n-1}+\eta_n. \tag{3.2c}$$

In the "static" version of the model, X_{n-1} contains only a constant, while two lags of π and U are added to X_{n-1} in the "dynamic" version. W_{1n} and W_{2n} are zero mean, unit–variance independent Gaussian shocks. Vector $\gamma = \left(\gamma_1, \gamma_{-1}^T\right)^T$ represents a government's beliefs about the Phillips curve; it is 6–dimensional in the "dynamic" and 2–dimensional in the "static" model. η_n is perceived by the government as white noise uncorrelated with regressors π_n and X_{n-1}.

The equilibrium is defined as a vector of beliefs $\overline{\gamma}$ at which the government's assumptions about orthogonality of η_n to the space of regressors are consistent with observations:

$$E\left[\eta_n\cdot\left(\pi_n, X_{n-1}\right)^T\right]=0. \tag{3.3}$$

CWS call this point a *self–confirming equilibrium*, or SCE. Williams [64] shows that at the SCE, $\gamma = \left(-\chi, 0, 0, 0, 0, u(1+\chi^2)\right)^T$, and the average inflation is $x_n = \chi u$. For a detailed description of the model, see CWS.

In a period n, the government solves (3.1), subject to (3.2b) and (3.2c), assuming that current beliefs γ_n will never change. The monetary policy action x_n is correctly anticipated by the private sector. U_n is generated according to (3.2a), and the government's beliefs are adjusted in a constant gain adaptive learning step. Let $\xi_n = \begin{bmatrix} W_{1n} & W_{2n} & X_{n-1}^T \end{bmatrix}^T$; $g(\gamma_n, \xi_n) = \eta_n \cdot \left(\pi_n, X_{n-1}^T\right)^T$; and $M_n(\gamma_n, \xi_n) = \left(\pi_n, X_{n-1}^T\right)^T \cdot \left(\pi_n, X_{n-1}^T\right)$. The next period's beliefs γ_{n+1} and R_{n+1} are given by

$$\gamma_{n+1} = \gamma_n + \epsilon R_n^{-1} g(\gamma_n, \xi_n), \quad (3.4a)$$

$$R_{n+1} = R_n + \epsilon \left(M_n(\gamma_n, \xi_n) - R_n\right), \quad (3.4b)$$

under RLS learning and by

$$\gamma_{n+1} = \gamma_n + \epsilon g(\gamma_n, \xi_n) \quad (3.5)$$

under the SG learning.[1]

Set the parameter vector $\theta_n^{\epsilon,SG}$ equal to γ_n for the SG and $\theta_n^{\epsilon,RLS} = \begin{bmatrix} \gamma_n^T, & vech^T(R_n) \end{bmatrix}^T$ for the RLS case.[2] Define $H^{RLS}(\theta_n^\epsilon, \xi_n) = \begin{bmatrix} \left(R_n^{-1} \cdot g(\gamma_n, \xi_n)\right)^T, & vech^T\left(M_n(\gamma_n, \xi_n) - R_n\right) \end{bmatrix}^T$ and $H^{SG}(\theta_n^\epsilon, \xi_n) = g(\gamma_n, \xi_n)$ to write the Stochastic Recursive Algorithm (SRA) in the standard form:

$$\theta_{n+1}^{\epsilon,j} = \theta_n^{\epsilon,j} + \epsilon H^j(\theta_n^{\epsilon,j}, \xi_n),\ j = \{RLS, SG\}, \quad (3.6a)$$

$$\xi_{n+1} = A(\gamma_n)\xi_n + B \cdot \begin{bmatrix} W_{1n+1} & W_{2n+1} \end{bmatrix}^T. \quad (3.6b)$$

Finally, the approximating ordinary differential equations corresponding to the above SRA are given by

$$\dot{\theta}^j = E[H^j(\theta^{\epsilon,j}, \xi_n)]. \quad (3.7)$$

The SCE (vector $\overline{\gamma}$ and corresponding 2^{nd} moments matrix $\overline{R}$ if RLS is used) is the only equilibrium of the above ODE. The SCE is stable for both RLS and SG in the dynamic and static versions of the model: the SCE is E–stable under both algorithms. The solution of the ODE (3.7) is called the "mean dynamics trajectory" of the SRA (3.6),

[1] R_n is the current estimate of the 2^{nd} moments matrix of the regressors.

[2] Following the notation of Lütkepohl [54], *vech* denotes a column vector in which abridged columns (the main diagonal and below) of a symmetric square matrix are stacked.

with the right–hand side of (3.7) being the "mean dynamics." For details and derivations, see Evans and Honkapohja [29]. Another local continuous–time approximation of the SRA around the SCE $\overline{\theta}$ can be derived in the constant gain case, as shown by Evans and Honkapohja [29, Prop. 7.8] and Williams [64, Theorem 3.2];

$$d\varphi_t^{RLS} = P\varphi_t dt + \sqrt{\epsilon}\Sigma^{1/2}(\bar{\theta}^{RLS})dW_t, \tag{3.8}$$

where $\varphi_t = \theta_t^{RLS} - \overline{\theta}^{RLS}$ are deviations from the SCE, see Kolyuzhnov, Bogomolova, and Slobodyan [50]. We use the approximation (3.8) to study behavior of the model when RLS learning is employed.

Still another variant of the mean dynamics approximation is the following difference equation obtained from (3.6a):

$$\theta_{n+1}^{\epsilon,j} = \theta_n^{\epsilon,j} + \epsilon \cdot E[H^j(\theta_n^{\epsilon,j}, \xi_n)]. \tag{3.9}$$

The difference between the above approximation and (3.7) is that ϵ is not assumed to be approaching zero asymptotically. This approximation turns out to be useful when we consider the learning dynamics in the SG case.

3.3 Behavior of Simulations

The discussion below refers to the model as parametrized in CWS: $\sigma_1 = \sigma_2 = 0.3$, $u = 5$, $\chi = 1$, $\beta = 0.98$.

3.3.1 Recursive Least Squares

Dynamic Model

It is well known that under the constant gain RLS learning, beliefs in the Phelps problem can exhibit "escapes": After a number of periods spent in the neighborhood of the SCE, the beliefs vector γ suddenly deviates from the SCE towards the "induction hypothesis" plane $\gamma_1 + \gamma_4 + \gamma_5 = 0$ ($\gamma_1 = 0$ axis for the static model) (see CWS, in particular Figures 6 and 7). During such an escape, the inflation rate falls from its Nash equilibrium value equal to χu and approaches 0 (see Figure 1 in CWS).

In Kolyuzhnov, Bogomolova, and Slobodyan [50], we have studied these escapes extensively and described the following sequence of events. If the constant gain parameter ϵ is not too small, the behavior of equation (3.4a) is almost one–dimensional because

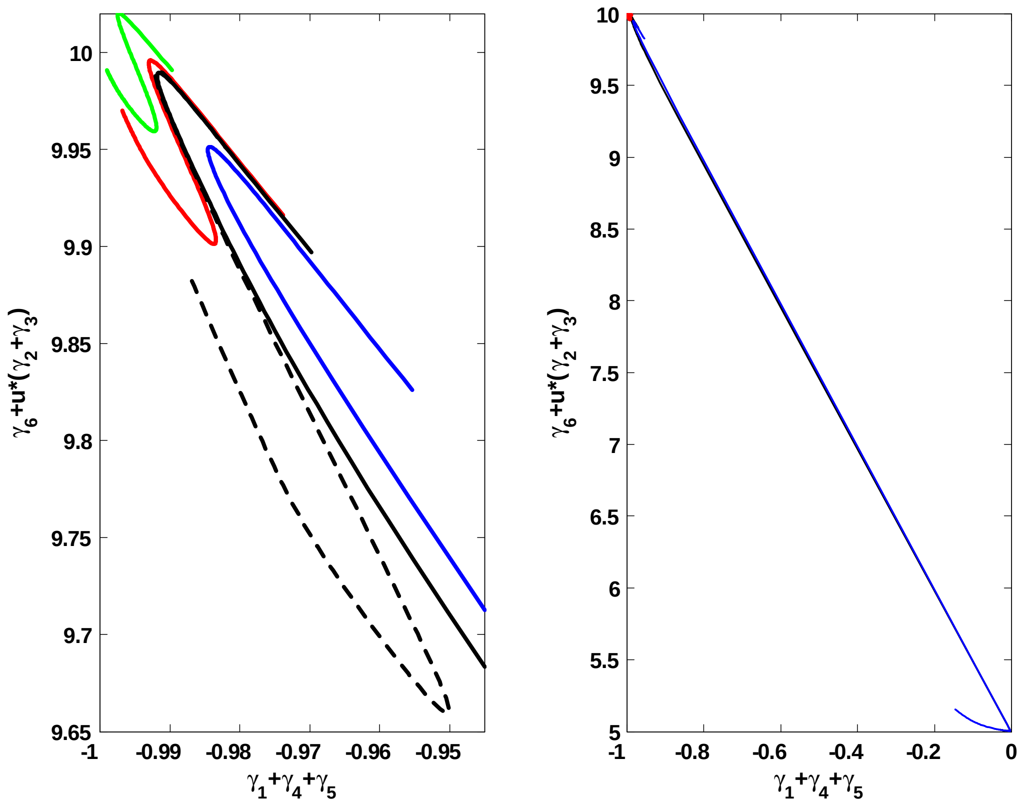

Figure 3.1: The mean dynamics trajectories under RLS.

the two largest eigenvalues of $\overline{R}^{-1}$, λ_1 and λ_2, equal 3083.8 and 29.1. As a result, the projection of $g(\gamma_n, \xi_n)$ onto v_1, the dominant eigenvector of $\overline{R}^{-1}$, is amplified about 100 times as strongly as the projection onto the second largest eigenvector. It is also well known that in this model, the region of attraction of the SCE is very small (see Figure 3.1 reprinted from Kolyuzhnov, Bogomolova, and Slobodyan [50] or Figures 8 and 9 in CWS). Outside of the immediate neighborhood of the SCE, the mean dynamics point away from it and towards the "induction hypothesis" plane in the direction which is very close to v_1. These trajectories linger in the neighborhood of the plane for a relatively long time and then start a slow return to the SCE. As a result, simulation runs with escapes tend to contain a set of points aligned along the dominant eigenvector of $\overline{R}^{-1}$ all the way towards the "induction hypothesis" plane, which is clearly demonstrated in the Figure 3.2, reprinted from Kolyuzhnov, Bogomolova, and Slobodyan [50].[3]

[3] In Figure 3.2, a 6–dimensional vector of beliefs γ is presented in the space of $(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$, defined as $\gamma_1 + \gamma_4 + \gamma_5$ and $u \cdot (\gamma_2 + \gamma_3) + \gamma_6$. A government's beliefs about the influence of past and current inflation on U_n are given by $\widetilde{\gamma}_1$, while $\widetilde{\gamma}_2$ represents the beliefs about the effect of past unemployment (and a

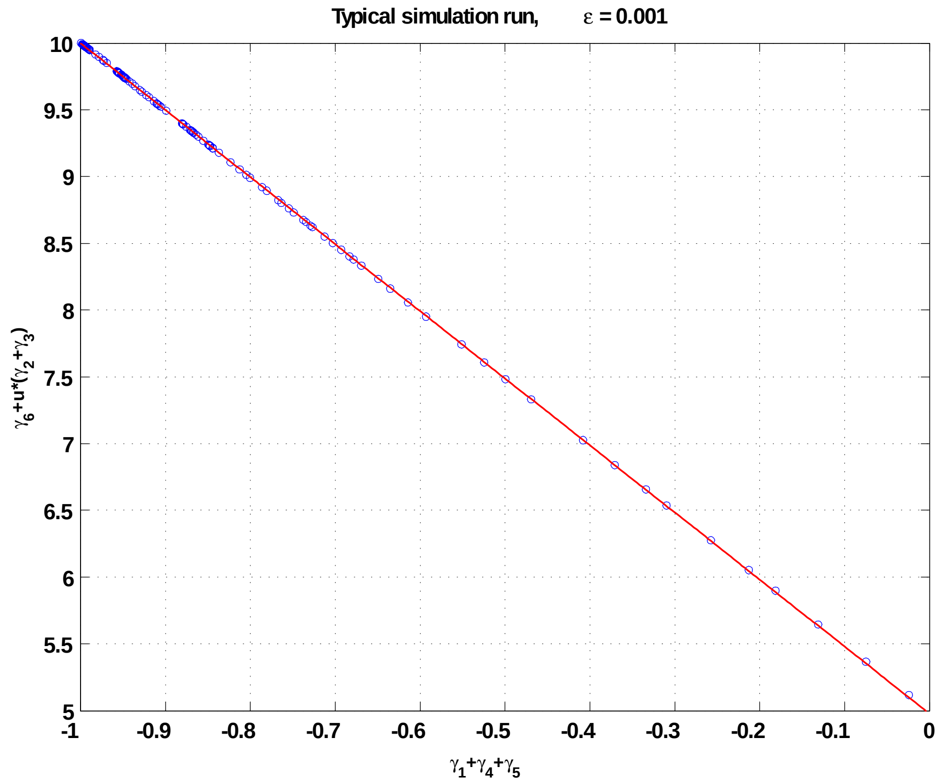

Figure 3.2: Typical simulation run and the "largest" eigenvector of R^{-1} under RLS.

We use this essential one–dimensionality to derive the following approximation of (3.8). Write $\varphi_t \approx x_t \cdot \widetilde{v}_1$, and multiply (3.8) by $\widetilde{v}_1^T$ from the left. The resulting 1–dimensional approximation is then given by

$$dx_t \approx \widetilde{v}_1^T D_\theta p(\bar{\theta}^{RLS})\widetilde{v}_1 \cdot x_t dt + \sqrt{\epsilon \widetilde{\lambda}_1} \cdot \widetilde{v}_1^T \cdot dW_t = A \cdot x_t dt + \sqrt{\epsilon \widetilde{\lambda}_1} dW_t, \tag{3.10}$$

where $\widetilde{\lambda}_1$ is the dominant eigenvalue of Σ. Note that $\widetilde{v}_1^T \cdot dW_t$ is a one–dimensional standard Brownian motion. (3.10) is then an Ornstein–Uhlenbeck process with well–known properties. In particular, one could easily derive the expected time until the process leaves any interval of the real line (see Borodin and Salminen [7]).[4]

To estimate the region of applicability of the approximation (3.10), take x_t^2 as

constant). The significant disbalance of eigenvalues of $\overline{R}^{-1}$ is inherited by the matrix Σ in (3.8), and the eigenvector v_1 is essentially collinear to the first 6 components of $\widetilde{v}_1$, the dominant eigenvector of Σ.

[4] Ornstein–Uhlenbeck approximation could also be useful in case one is interested in selecting the value of ϵ such that for a given time period the probability of observing an escape is below some given threshold (dynamics under learning is empirically stable).

the Lyapunov function and calculate LV for one–dimensional diffusion (3.10):[5]

$$LV = 2 \cdot \left(A x_t^2 + \epsilon \widetilde{\lambda}_1 \right).$$

Clearly, LV is positive for small x_t, and thus $V(x_t) = x_t^2$ is expected to *increase.* In other words, in a small neighborhood of the SCE, the Stochastic Recursive Algorithm (3.6) is expected to be locally divergent on average. We would call values of ϵ "small" if for x_t corresponding to the boundary of the SCE's stability region under the mean dynamics, the value of LV is negative: Once the SRA approaches this boundary, it is expected to turn back towards the SCE. If such behavior is observed, one expects the invariant distribution derived along the lines of Evans and Honkapohja [29, Ch. 14.4] to be valid, and other methods of describing escape dynamics are needed, such as the Large Deviations Theory (see CWS and Kolyuzhnov, Bogomolova, and Slobodyan [50]). For values of ϵ which are not "small," the approximation (3.10) could be used to derive expected escape time. In the dynamic model, values of ϵ below $2 \cdot 10^{-5}$ are "small."

What is the right ϵ and the time scale?

How should one approach the problem of choosing ϵ? Putting aside any considerations related to the stability of learning in a particular model, two rules of thumb for selecting ϵ seem sensible. The first is based on the fact that constant gain adaptive learning is well suited to situations with time–varying parameters or structural breaks. In this case, $1/\epsilon$ should be related to the typical time which is needed to observe a break, or for the time variation to become "significant." Alternatively, one could imagine that the initial value of parameters is obtained through some method of statistical estimation such as OLS. In this case, it is natural to assign to every point in the initial estimation a weight equal to $1/N$. If there is no reason to believe that subsequent points are in some sense superior to those used to derive an initial estimate, the constant gain ϵ should be comparable to $1/N$. Given the nature of the Phelps problem where inflation might be available on a monthly basis but the output gap could be evaluated only quarterly, values of ϵ not much larger or smaller than 0.01 seem empirically justified. In a recent paper, Orphanides [58] considers values of ϵ between 0.01 and 0.03 as fitting the data in a model

[5]The operator L defined for a function V has the following meaning: Under certain conditions, the expected value of $V(t, X(t)) - V(s, X(s))$ is given as an integral from s to t over LV (see Khasminskii [48, Ch. 3]). In some sense, in stochastic differential equations LV plays the role of a time derivative of the Lyapunov function $\frac{dV}{dt}$ for the deterministic system.

	Dynamic model		Static model	
ϵ	Simulations	Theory	Simulations	Theory
$2 \cdot 10^{-5}$	$1.10 \cdot 10^5$	$1.86 \cdot 10^5$		
$3 \cdot 10^{-5}$	$5.10 \cdot 10^4$	$7.21 \cdot 10^4$	$4.40 \cdot 10^7$	$9.40 \cdot 10^8$
$5 \cdot 10^{-5}$	$1.88 \cdot 10^4$	$2.34 \cdot 10^4$	$1.93 \cdot 10^6$	$9.90 \cdot 10^6$
$1 \cdot 10^{-4}$	$4.84 \cdot 10^3$	$5.43 \cdot 10^3$	$1.50 \cdot 10^5$	$2.75 \cdot 10^5$
$2 \cdot 10^{-4}$	$1.26 \cdot 10^3$	$1.31 \cdot 10^3$	$2.38 \cdot 10^4$	$2.97 \cdot 10^4$
$4 \cdot 10^{-4}$	336.96	321.5	$5.06 \cdot 10^3$	$5.26 \cdot 10^3$
$1 \cdot 10^{-3}$	64.59	50.9	733.57	701.5
$2 \cdot 10^{-3}$	21.49	12.68	189.98	165.7
$3 \cdot 10^{-3}$	12.50	5.63	87.00	72.27
$4 \cdot 10^{-3}$	8.77	3.16	52.08	40.28
$5 \cdot 10^{-3}$	6.79	2.02	34.39	25.64
$6 \cdot 10^{-3}$	5.99	1.40	24.76	17.74
$7 \cdot 10^{-3}$	4.98	1.03	19.14	13.00
$8 \cdot 10^{-3}$	4.49	0.79	15.02	9.93
$9 \cdot 10^{-3}$	4.12	0.62	13.32	7.84
$1 \cdot 10^{-2}$	3.70	0.51	11.16	6.34

Table 3.1: A comparison of the theoretically derived values of expected escape time and empirically observed average escape times.

with constant gain RLS learning. He also uses $\epsilon = 0.005$ for the SG constant gain learning of a natural real rate and a natural unemployment rate.

Notice that the period in the Phelps model could not be shorter than a quarter (or a month). As Table 3.1 shows, for $\epsilon < 1 \cdot 10^{-4}$ in the dynamic model and $\epsilon < 4 \cdot 10^{-4}$ in the static one, the expected time until escape becomes larger than an economically relevant time scale (say, a hundred years); probability of observing an escape within this time becomes negligible as ϵ decreases even further. An important caveat to this statement is that both the theoretical and simulation results are obtained by imposing the SCE as the starting point of learning. In other words, one starts from a situation of completed learning, where the government and the private sector are playing Nash equilibrium, and is interested in the expected time until the economy "unlearns" the Nash equilibrium given a particular constant gain learning rule. If, instead of the SCE, initial beliefs are given by a point which is closer to the stability region's boundary, one would expect smaller escape times.

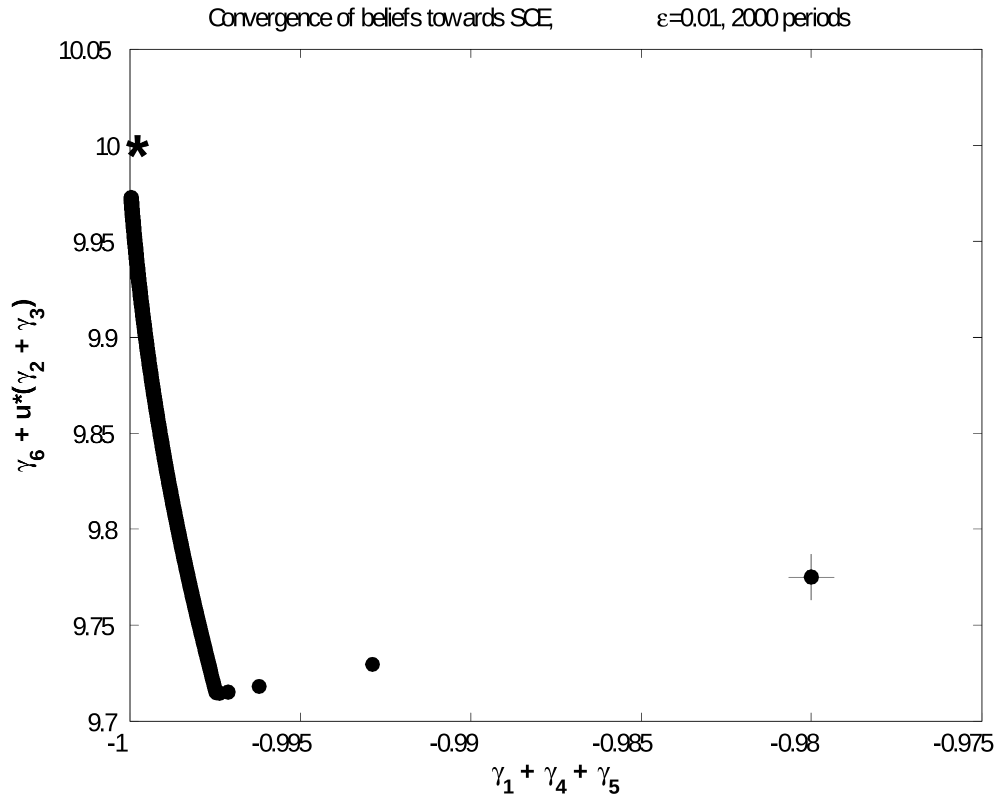

Figure 3.3: Iterations of the mean dynamics map. '+' sign — start of the simulation. '*' — the SCE location.

Static Model

Dynamics of the static model under the constant gain RLS learning is qualitatively similar to that of the dynamic one: a move out of the immediate region of attraction of the SCE, followed by a long trek to the Ramsey equilibrium outcome with zero average inflation. The dynamics is essentially one–dimensional. However, the radius of the region of attraction is slightly larger in the dominant direction than in the dynamic model, and the diffusion is less powerful.[6] As a result, in the static model ϵ starts to be "small" at about $3 \cdot 10^{-4}$.

The combined effect of the stronger drift, weaker diffusion, and larger stability region is obvious: a significantly larger than in the dynamic model expected number of periods until the simulations escape the neighborhood of the SCE. Table 3.1 compares empirically observed average time needed to escape with the theoretically predicted values

[6] In the dynamic (static) model, A=-0.41 (-0.52) and $\widetilde{\lambda}_1$ =278 (26).

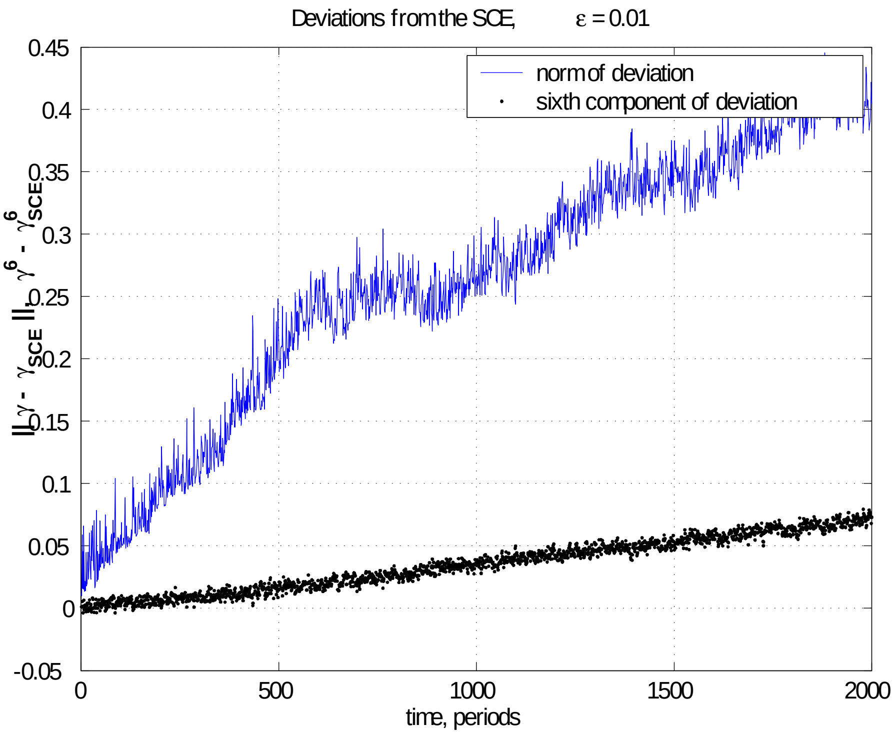

Figure 3.4: Divergence of a simulation run from the SCE.

for different choices of the constant gain parameter ϵ. For values of ϵ which are not "small," the agreement is rather good, especially for the static model. In agreement with our estimate of the Ornstein–Uhlenbeck approximation's applicability, it starts to overpredict for "small" ϵ. This effect is especially pronounced for the static model.

3.3.2 Stochastic Gradient Learning

It is necessary to note that in the SG case, the dependence of the learning dynamics on ϵ is dramatically different from the RLS case. In a nutshell, simulations are divergent for a rather wide interval of ϵ. On the other hand, the term R_n^{-1} does not multiply the right–hand–side in Eq. (3.5), which prevents usage of a one–dimensional approximation which proved to be so successful in the RLS case.

Dynamic Model

In the approximation (3.9), the matrix

$$F(\epsilon) = I + \epsilon D_\theta p(\bar{\theta}^{SG})$$

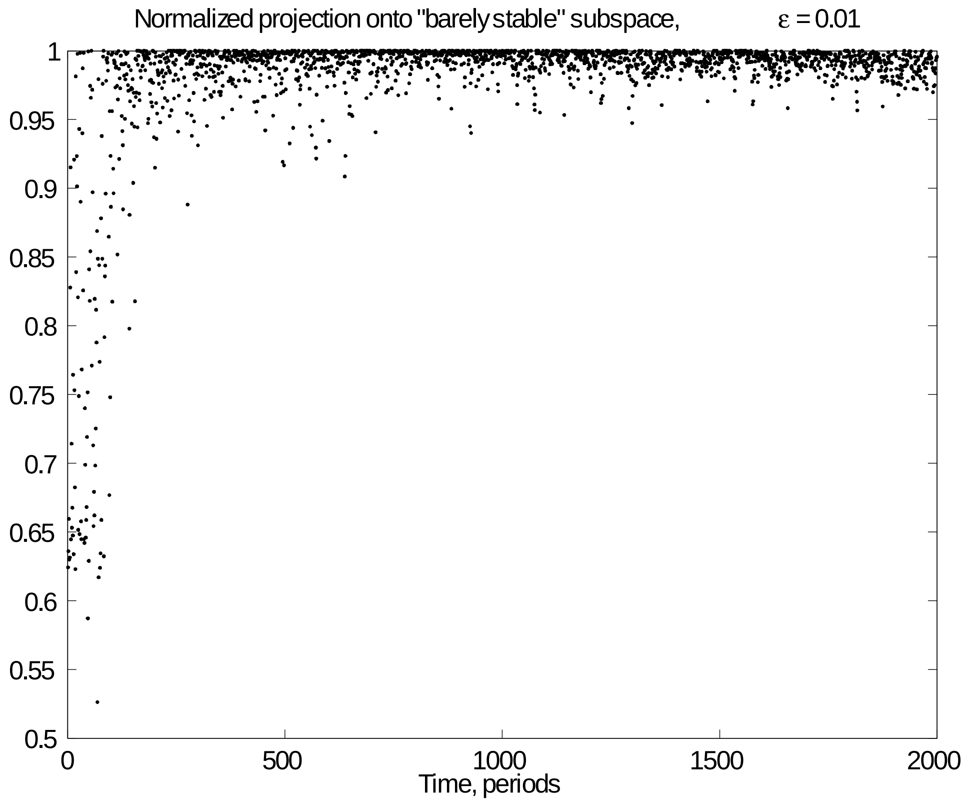

Figure 3.5: Projection of beliefs onto the subspace spanned by the mean dynamics map's almost unitary eigenvalues' eigenvectors in a typical simulation run.

is stable but only just: For $\epsilon = 0.01$, its eigenvalues range from λ_1 =0.2447 to λ_2 =0.9988 to λ_6 =0.99999862. Five out of six eigenvalues are almost unitary. Under the mean dynamics (3.9), any deviation from the SCE results in a fast movement along x_1, the eigenvector which corresponds to λ_1, and then an extremely slow convergence back to the SCE along the remaining five directions, see Figure 3.3. On the other hand, simulations of (3.6) behave very differently. Figure 3.4 plots a norm of deviations from the SCE and $\gamma_6 - \overline{\gamma}_6$: There is a clearly distinguishable movement away from the SCE which seems almost deterministic.[7] For this value of ϵ, the inflation rate will drop below 4 (at the SCE, mean inflation equals 5) in a couple of hundred periods, which is definitely the time scale with which one should be concerned. How could one explain the discrepancy between the mean dynamics (3.9) and the simulations?

Figure 3.5 plots a projection of $\frac{\gamma_n - \overline{\gamma}}{\|\gamma_n - \overline{\gamma}\|}$ onto the sub–space spanned by five eigen-

[7]If we observe the simulations for a larger number of periods, the belief vector γ eventually reaches values at which the state vector process loses stationarity, and the simulation breaks down.

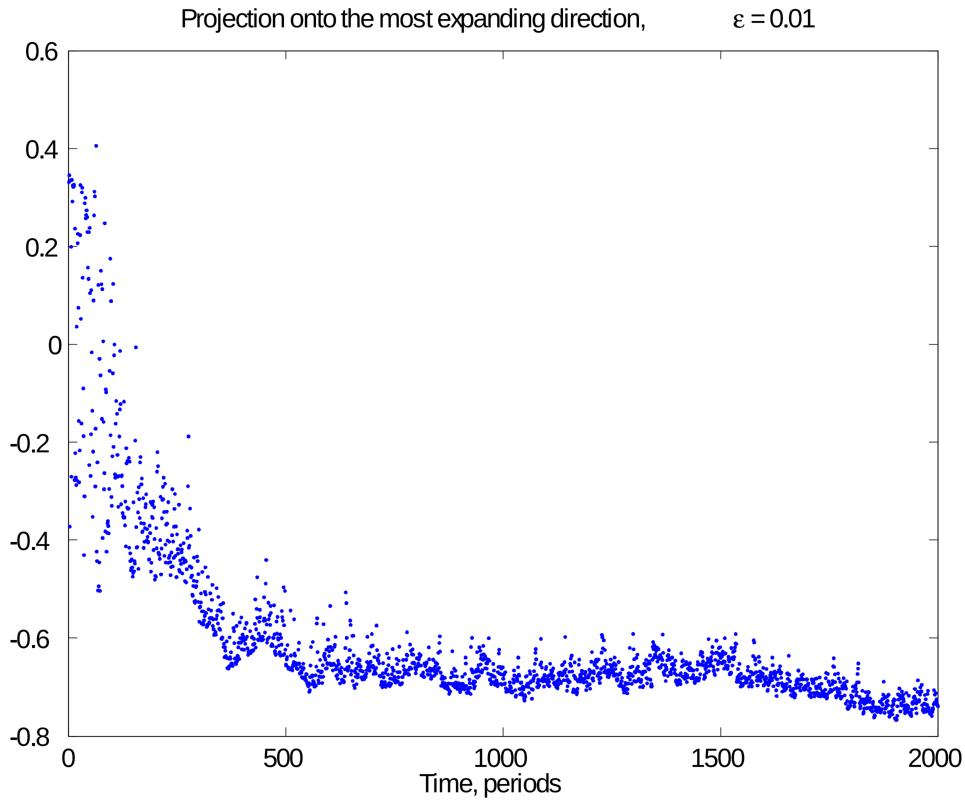

Figure 3.6: Projection of beliefs onto the expansive direction of the mean dynamics map in a typical simulation run.

vectors of $F(\epsilon)$ which correspond to the almost unitary eigenvalues for a typical simulation run with $\epsilon = 0.01$. Within the first hundred simulation periods, this projection becomes very close to unity: average value for the first ten (hundred) periods is 0.69 (0.80). Thus, a simulation run quickly approaches some neighborhood of the sub–space and does not leave it for any extended period of time. This behavior is natural: Any initial deviation along x_1 will shrink to $0.25^3 \sim 1.5\%$ of its initial size in just 3 steps. On the other hand, deviations along five other eigenvectors will take at least $\frac{\ln(0.5)}{\ln(0.9988)} \sim 577$ periods to reach 50% of their initial magnitude.

Another feature of the matrix $F(\epsilon)$ which helps to explain the behavior of simulations is the presence of directions along which deviations are expected to *increase* before declining. Such directions exist because the symmetric part of $F(\epsilon)$, $F^{sym}(\epsilon) = \frac{F(\epsilon)+F(\epsilon)^T}{2}$, is not stable. After one iteration of the map $F(\epsilon)$, initial deviation in the direction w, the unstable eigenvector of $F^{sym}(\epsilon)$, is expected to *increase* its projection onto w and thus to

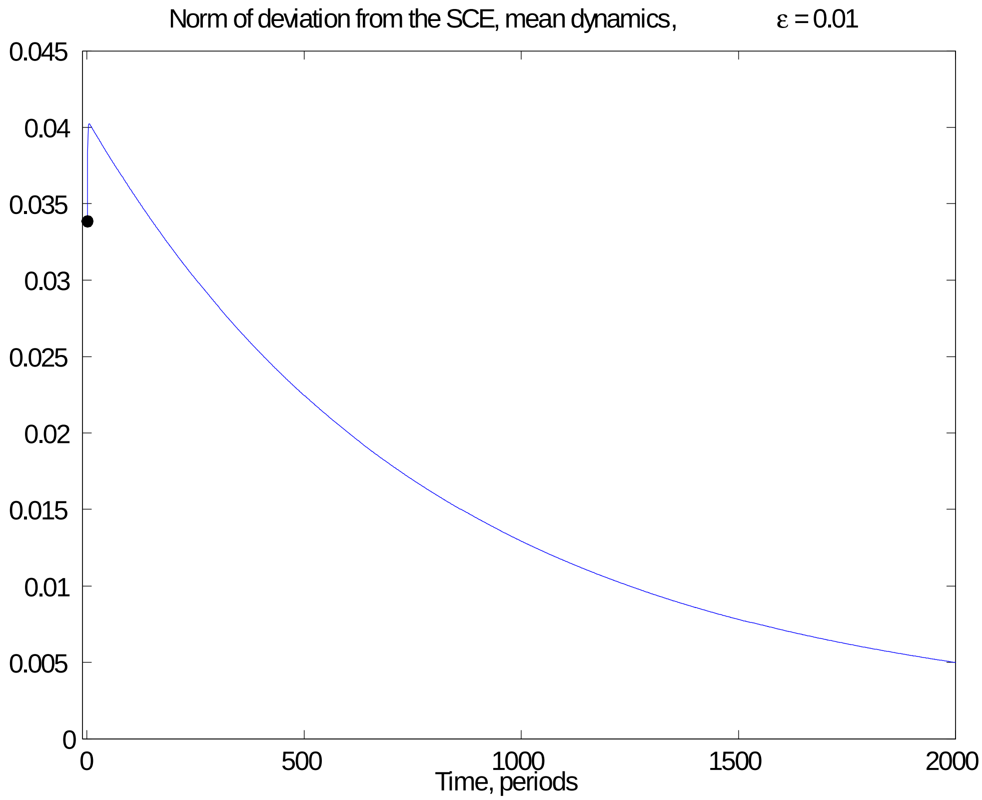

Figure 3.7: Evolution of beliefs under iterations of the mean dynamics map. Initial deviation in the expanding direction.

increase its norm.[8] The largest eigenvalue of $F^{sym}(\epsilon)$ equals 1.103 at $\epsilon = 0.01$, 1.01 at $\epsilon = 0.001$ and 1.001 at $\epsilon = 1 \cdot 10^{-4}$. A projection of $\frac{\gamma_n - \overline{\gamma}}{\|\gamma_n - \overline{\gamma}\|}$ onto w is plotted in Figure 3.6 (only the absolute value of the projection matters, not its sign). It becomes large very fast, in about one hundred simulation periods or less. The system (3.9) is expected to demonstrate a locally divergent behavior whenever this projection is large. To support further the crucial importance of the projection onto w, Figure 3.7 presents the norm of deviation from the SCE for the mean dynamics trajectory which started from a point γ that lies in the direction w. There is a steep initial increase in the norm, followed by a long decline which is still far from complete after 2000 periods. To overcome the initial increase and return the system to the norm of deviation equal to its initial value, 150 periods are needed.

[8] Suppose an initial deviation is given by w. After one period, this deviation is transformed into Fw. Projection $w^T Fw$ then gives a measure of expansion or contraction in the direction of w after one iteration of map F. But for any vector w, $w^T Fw = w^T F^{sym} w$. Therefore, in order to find expanding (after one iteration) directions of F, one could look at eigenvalues of F^{sym}.

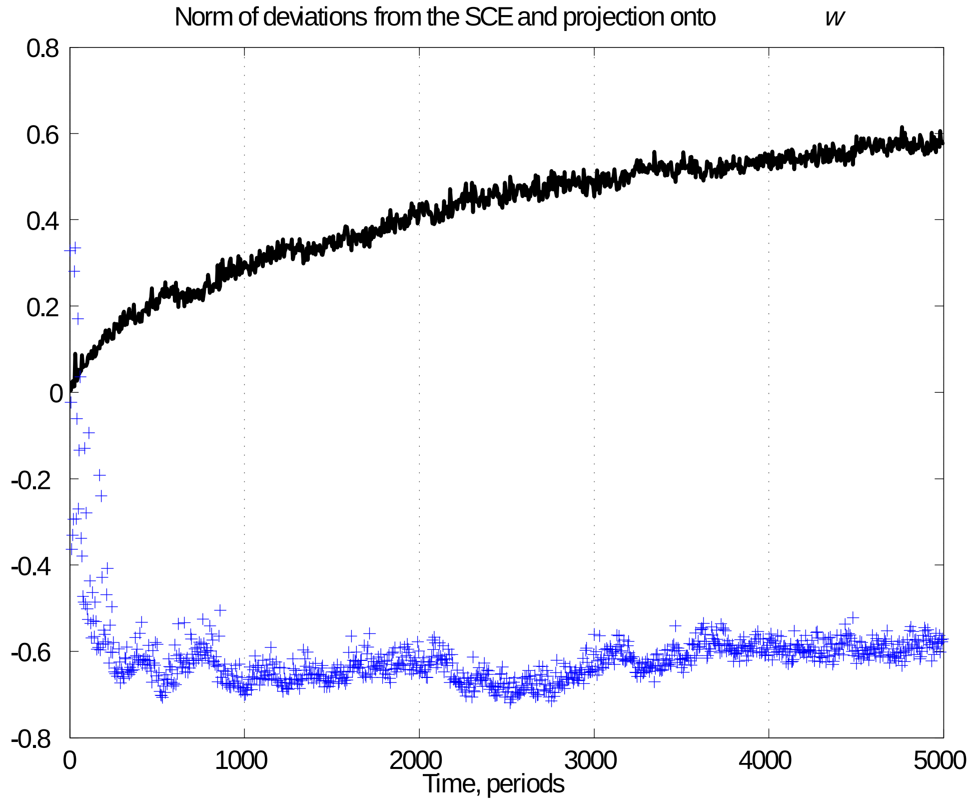

Figure 3.8: The largest is the projection of beliefs onto the expanding direction of the mean dynamics map ('+' sign), the faster the beliefs deviate from the SCE ('.' sign).

The norm of the projection of w onto the sub–space spanned by the five eigenvectors is rather large and equals 0.95. When the dynamics of (3.6) is restricted almost exclusively to this subspace, mean dynamics plays almost no role in the short run. Random disturbances are then very likely to produce the value of $\gamma_n - \overline{\gamma}$, which has a significant projection onto w during the 150 periods which are needed to eliminate the effect of the previous shock in this direction. Once such shock happens, the projection is not likely to disappear given a very weak stabilizing force of the mean dynamics on the sub–space.

As a final piece of evidence connecting the vector w with the divergent behavior of simulations, consider Figure 3.8. In the periods when the projection of $\frac{\gamma_n - \overline{\gamma}}{\|\gamma_n - \overline{\gamma}\|}$ onto w (crosses) is particularly large, the distance between the beliefs γ_n and the SCE $\overline{\gamma}$ (solid line) grows the fastest; a relative decline in the projection is correlated with a temporary stop or even a reversal of the divergent behavior.

Summarizing the discussion, we could say that a clear instability observed in the behavior of the SRA for SG learning in the dynamic Phelps problem is caused by a particular structure of the mean dynamics map $F(\epsilon)$. The sub–space spanned by the almost unitary eigenvalues' eigenvectors of $F(\epsilon)$ is almost parallel to the direction along which the mean dynamics is expanding in the short run rather than contracting. Given that any random deviation away from the subspace is likely to be very short–lived, and that a contracting mean dynamics within the sub–space is very weak, random vectors with a relatively large projection onto the expansive direction are likely to appear. Once such a projection appears, it is unlikely to be averaged away by the mean dynamics.

We checked the behavior of the algorithm for other values of ϵ. Qualitatively, the picture does not change: There is still an apparent divergence of the vector of a government's beliefs γ_n away from the SCE. One could still observe a very fast convergence towards the sub–space spanned by the five almost unitary eigenvalues' eigenvectors and a significant projection onto the expanding direction w. Only for very small values of $\epsilon \leq 8 \cdot 10^{-6}$ we start observing a different behavior, when the system (3.6) does not systematically diverge and fluctuates in some neighborhood of the SCE.

Static Model

Taking into account that under RLS learning the static model was much more stable (it took much longer for the escape to the "induction hypothesis" plane to happen), we expect this feature to be preserved under SG learning as well. This is what is indeed observed. Clearly unstable behavior is observed only for relatively large values of ϵ above $3 \cdot 10^{-2}$. This instability could take two forms: either a convergence to a quasi–stable stochastic steady state where $\|\gamma - \overline{\gamma}\|$ is about 3 for ϵ between approximately $6.5 \cdot 10^{-2}$ and $7.9 \cdot 10^{-2}$ (above $\epsilon \sim 7.9 \cdot 10^{-2}$, the mean dynamics map $F(\epsilon)$ has a real eigenvalue which is less than -1 making the SCE unstable), or a divergence of simulations from the SCE for $3.5 \cdot 10^{-2} \lesssim \epsilon \lesssim 6.5 \cdot 10^{-2}$. When ϵ equals $3.5 \cdot 10^{-2}$ or less, empirically relevant time scales are characterized by what seems to be stable dynamics. The speed of divergence significantly depends on the value of ϵ: While at $\epsilon = 5 \cdot 10^{-2}$, less than 100 iterations are typically needed to observe a deviation from the SCE such that $\|\gamma - \overline{\gamma}\| \geq 0.1$; such large excursions are not likely to be observed before the 500^{th} iteration for $\epsilon = 4 \cdot 10^{-2}$. As in the dynamic model, the eventual outcome of divergent simulations is the value of γ which leads to at least one eigenvalue of the matrix $A(\gamma)$ in (3.6b) being outside of the unit circle and thus to a non–stationary state process.

Applying the reasoning demonstrated above to the dynamics of the static model under SG learning in real time, we could say the following. The map $\mathcal{F}(\epsilon)$ has two eigenvalues. One is always close to one (0.9999 for $\epsilon = 3 \cdot 10^{-2}$). The other is a linearly decreasing function of ϵ. It equals -1 when $\epsilon \sim 7.9 \cdot 10^{-2}$ and approaches 1 as $\epsilon \to 0$. It is still true that the divergent behavior is related to the movement along the almost unitary eigenvalue's eigenvector: Projection of w onto this eigenvector equals 0.9988, and the fastest divergence of beliefs from their SCE values occurs when $\gamma - \overline{\gamma}$ is in the closest alignment with w ($w^T \cdot \frac{\gamma_n - \overline{\gamma}}{\|\gamma_n - \overline{\gamma}\|}$ is close to one). There are two crucial differences with the dynamics model, however. First, the direction w is very weakly expansive: The unstable eigenvalue of $\mathcal{F}^{sym}(\epsilon)$ equals only 1.0018 when $\epsilon \sim 3 \cdot 10^{-2}$ and becomes even smaller as ϵ decreases. At the same time, the dominant eigenvalue of $\mathcal{F}(\epsilon)$ equals 0.23 for $\epsilon \sim 3 \cdot 10^{-2}$ and is decreasing in ϵ. Thus, for smaller values of ϵ, the dynamics of (3.9) loses its essentially one–dimensional nature in the expanding direction, and the expansive movement in the direction w is not too strong (compare 1.0018 to the 1.103 reported for the dynamic model). Instead of 150 periods needed to start reversing a deviation in the direction of w, which we reported for the dynamic model at $\epsilon = 0.01$, only 3-4 iterations are needed to achieve the same result in the static model at similar values of ϵ. It is not a big surprise, then, that the static model under the SG learning stops diverging at much larger values of the constant gain.

Reasons for difference with the RLS case

Why do we observe the diverging behavior documented above only in the SG case? RLS case differs from the SG one in three respects. First, the mean dynamics is very weak relative to the stochastic dynamics especially in the direction of dominant eigenvector of $\overline{R}^{-1}$, as documented in Kolyuzhnov, Bogomolova, and Slobodyan [50]. Second, the mean dynamics map $\mathcal{F}(\epsilon)$ does not contain strongly contracting eigenvalues. And third, those eigenvalues of $\mathcal{F}(\epsilon)$ that are closest to the unit circle are much further from it than in the SG case. A combination of these three factors assures that even though simulation runs under RLS learning do exhibit relatively large projections in the expanding direction of $\mathcal{F}^{sym}(\epsilon)$, these projections are not correlated with episodes of particularly fast deviations from the SCE.

3.4 Conclusion

We compared the performance of two methods of adaptive learning with constant gain, Recursive Least Squares and Stochastic Gradient learning, in a Phelps model of a monetary policy which has been extensively studied previously. For the values of ϵ which might be justified for the problem, it is a well–known fact that the RLS adaptive learning could force the government's beliefs about the Phillips curve to "escape," or deviate significantly, from the neighborhood of the Self–Confirming Equilibrium, where the inflation level is set at high levels, towards the beliefs which lead the policy maker to set inflation close to zero. We approximated the discrete–time Stochastic Recursive Algorithm which describes RLS constant gain learning by a one–dimensional continuous–time Ornstein–Uhlenbeck process and derived expected escape times out of a small neighborhood of the SCE. The theoretical prediction works rather well when compared with the simulation results.

Turning our attention to the SG learning, we showed that the model dynamics is divergent for a large interval of values of ϵ. The divergence is especially pronounced when SG learning is used in the dynamic version of the Phelps problem. This behavior is caused by the existence of eigenvalues of the SRA mean dynamics map which are very close to the unit circle; deviations in the direction of corresponding eigenvectors contract very slowly. Moreover, the SRA mean dynamics map has direction where deviation is expected to expand in the short run rather than contract, and this direction is almost parallel to the sub–space spanned by the slowly contracting eigenvectors. Such a combination leads to a divergent behavior of the SRA, which is reversed only for the very small ϵ values when the expansion rate becomes very small. Behavior of the static model exhibits similar features, with a crucial difference of the expansion rate: For the empirically relevant values of ϵ, it is less than 1.02 instead of 1.1 as in the dynamic model. This difference means that the SRA stops exhibiting divergent behavior for much larger values of the constant gain parameter in the static than in the dynamic model.

Comparing the two variants of the model under two types of constant gain adaptive learning, we could say that only SG learning in the static model demonstrates an absence of large excursions of beliefs from the SCE at an empirically relevant time scale and for constant gain values likely to be used in practice ("stability"). Additionally, the expected escape time rises very steeply as ϵ decreases. Following Evans, Honkapohja, and

Williams [32], one could thus endorse using this adaptive learning method for the static model. The overall result, however, cannot be judged as very good as three out of four modifications produce an “unstable” result.

A very unbalanced nature (large differences between the dominant eigenvalue and the rest) of the second moments matrix $\overline{R}$ plays a significant role in the results, making the stochastic dynamics strongly one–dimensional in the RLS case and leading to almost unitary eigenvalues in the SG case. Whether this feature is caused by the fact that the government uses a mis–specified model in the Phelps problem warrants further investigation.

The behavior of the SRA under SG learning in real time leads us to express a warning. Checking that the mean dynamics map is asymptotically stable is not enough to guarantee “stable” behavior of the constant gain learning algorithm in real time; moreover, checking that the mean dynamics trajectories are stable in a large region is not enough either. If many eigenvalues of the mean dynamics map for a constant gain learning algorithm are close to the unit circle, and the mean dynamics map is not contracting in every direction, the Stochastic Recursive Algorithm might exhibit *divergent* behavior despite *convergent* mean dynamics.

Part II

Heterogeneous Learning

Chapter 4

Economic Dynamics Under Heterogeneous Learning: Necessary and Sufficient Conditions for Stability

Economic Dynamics Under Heterogeneous Learning: Necessary and Sufficient Conditions for Stability*

Dmitri Kolyuzhnov†

CERGE–EI‡
Politických vězňů 7, 111 21 Praha 1,
Czech Republic

Abstract

I provide sufficient conditions and necessary conditions for stability of a structurally heterogeneous economy under heterogeneous learning of agents. These conditions are written in terms of the structural heterogeneity independent of heterogeneity in learning. I have found an easily interpretable unifying condition which is sufficient for convergence of an economy under mixed RLS/SG learning with different degrees of inertia towards a rational expectations equilibrium for a broad class of economic models and a criterion for such a convergence in the univariate case. The conditions are formulated using the concept of a subeconomy and a suitably defined aggregate economy. I demonstrate and provide interpretation of the derived conditions and the criterion on univariate and multivariate examples, including two specifications of the overlapping generations model and the model of simultaneous markets with structural heterogeneity.

JEL Classification: C62, D83, E10
Keywords: adaptive learning, stability of equilibrium, heterogeneous agents

---

*The major part of this paper was written while the author stayed as a visiting Ph.D. student at the University of Cambridge, UK. The author expresses special thanks to his supervisor while at Cambridge, Seppo Honkapohja, for continuous support of this research and immeasurable help and advice. All errors are the author's responsibility.

The author expresses thanks to the participants of the European Economics and Finance Society (EEFS 2006) 5th Annual Conference "European Labour Markets, Trade and Financial Flows and Economic Growth" in Heraklion, Crete, Greece, May, 18-21, 2006 and CEF 2006, Limassol, Cyprus, June, 17-21, 2006 for valuable discussion.

†*Dmitri.Kolyuzhnov@cerge-ei.cz.*

‡CERGE–EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic.

4.1 Introduction

Until some time ago, works studying models of economic dynamics assumed rational expectations of agents. However, the need to study models under bounded rationality of agents was well argumented in Sargent [60]. Later this approach was also adopted (among others) in works of Evans and Honkapohja, and a standard argument in defense of bounded rationality can be found in Evans and Honkapohja [29], as well as in Sargent [60].

The rational expectations (RE) approach implies that agents have a lot of knowledge about the economy (e.g., of the model structure and its parameter values). However, in empirical work, economists who assume RE equilibria in their theoretical model do not know the parameter values and must estimate them econometrically. According to the argument of Sargent [60], it appears more natural to assume that in a given economy agents face the same limitations. It is then suggested to view agents as econometricians when forecasting the future state of the economy. Each time agents obtain new observations, they update their forecast rules. This approach introduces a specific form of bounded rationality captured by the concept of adaptive learning.

The bounded rationality approach can serve several purposes, for example, to test the validity of the RE hypothesis by checking if a given dynamic model converges over time to the rational expectations equilibrium (REE) implied by the model (under the RE hypothesis), or for equilibrium selection (in models with multiple equilibria). In both cases we have to analyze convergence of our model under adaptive learning to a REE. To do this, we need to check certain stability conditions. This introduces the area of my research: studying stability conditions in models with adaptive learning.

Following the adaptive learning literature, I consider two possible algorithms used to reflect bounded rationality: the generalized recursive least squares (RLS) and the generalized stochastic gradient (SG) algorithms. Both algorithms are examples of econometric learning.[1] Each period agents update the parameter estimates in the following way: the updated parameter estimate equals the previous estimate plus a linear function of the most recent forecast error multiplied by the gain parameter, capturing how important is the forecast error to the agent. The description of both algorithms can be found,

[1] One more type of econometric learning is Bayesian learning. See Honkapohja and Mitra [43] for references of other forms of learning – like bounded memory rules and non-econometric learning (including computational intelligence algorithms).

for example, in Evans and Honkapohja [29]; Giannitsarou [37]; Evans, Honkapohja, and Williams [32]; and Honkapohja and Mitra [43].

The difference between the two algorithms is that the RLS algorithm[2] has two updating equations: one for updating the parameters entering the forecast functions, the other – for updating the estimates of the second moments matrix of these parameters. The SG algorithm assumes this matrix fixed (reflecting modeling of "less sophisticated" agents).

The first papers taking the bounded rationality approach of Sargent [60] considered an economy of a representative agent (assuming that all agents follow the same learning algorithm, be it RLS or SG). Later, some works began to introduce heterogeneity in the updating procedure. The idea was to check whether the representative agent hypothesis (implied by homogenous learning) in learning influences stability results. It has been demonstrated that, in general, stability under homogeneous learning does not imply stability under heterogeneous learning. Examples of such works are Giannitsarou [37], who assumed agents homogeneous in all respects but the way they learn, and Honkapohja and Mitra [43], who consider a general setup assuming both structural heterogeneity and heterogeneous learning. Both papers study stability conditions of the economy. Honkapohja and Mitra [43] derive a general stability criterion, in which stability is defined both in terms of the model structure and learning characteristics.

The difference in learning characteristics across agents means heterogeneity in learning. Among these learning characteristics are initial perceptions meaning that agents may have different perceptions about the economy reflected in different initial values in their learning algorithms; the type of the updating algorithm: RLS or SG (reflecting "sophisticated" and "less sophisticated" agents, respectively); and parameters of the updating algorithm (degree of inertia) — relative weights agents put on the most recent forecast error, while updating the parameter estimates in their forecast functions (it can be called the speed of updating, reflecting how agents differ in their reaction to innovation).

A combination of all differences in the learning characteristics described above can be expressed by the type of learning when one part of agents uses the RLS algorithm and the other part uses the SG algorithm, and all of them have different degrees of inertia

[2] The RLS algorithm (non-generalized) can be obtained from OLS estimation of parameters by rewriting it in the recursive form. The generalized RLS is derived from RLS by substituting the gain sequence $1/t$ used in updating the regression coefficients with any decreasing gain sequence.

as well as different initial perceptions. Such type of learning algorithm is called mixed RLS/SG learning with (possibly) different degrees of inertia.

In my paper I solve the following open question posed by Honkapohja and Mitra [43] — to find conditions for stability of a structurally heterogeneous economy under mixed RLS/SG learning with (possibly) different degrees of inertia in terms of structural heterogeneity only, independent of heterogeneity in learning.

Though Honkapohja and Mitra [43] have formulated a general criterion for such stability and have been able to solve for sufficient conditions for the case of a univariate model (model with one endogenous variable), they did not derive the conditions (necessary, and/or sufficient) in terms of the model structure only, independent of learning characteristics, for the general (multivariate) case with an arbitrary number of agent types and any degree of inertia.

As, in essence, the criterion (in its sufficiency part) for stability of a structurally heterogeneous economy under mixed RLS/SG learning by Honkapohja and Mitra [43] implies looking for sufficient conditions for D–stability of a particular stability Jacobian matrix corresponding to the model, where, according to Johnson [44, p. 54], "the D–stables are just those matrices which remain stable under any relative reweighting of the rows or columns," I use different sets of sufficient conditions for D–stability of this Jacobian matrix and simplify them using a particular structure of the model, and try to provide the derived conditions with some economic interpretation.

Specifically, in this paper I conduct a systematic analysis of this problem. First, I analyze what has been done so far in mathematics on deriving sufficient conditions for stability of a matrix in the most general setup of a matrix differential equation: $\dot{x} = Ax+b$, where A has the form $D\Omega$, with D being a positive diagonal matrix. It has turned out that *the most general* results can be grouped according to the perspective from which the problem was approached.

One group of results is based on the Lyapunov theorem[3] and its application to D–stability by the theorem of Arrow and McManus;[4] another group is based on the negative diagonal dominance condition which is sufficient for D–stability (McKenzie theorem[5]); a third set of results can be derived from the characteristic equation analysis,

[3]See Theorem B.2 in Appendix B.2.
[4]See Theorem B.3 in Appendix B.2.
[5]See Theorem B.4 in Appendix B.3.

using Routh–Hurwitz necessary and sufficient conditions[6] for negativity of all eigenvalues of the polynomial of order n; and the last set of sufficient results can be derived using an alternative definition of D–stability[7] that allows to bypass the Routh–Hurwitz conditions.

Among the approaches mentioned above, the ones that are based on the negative diagonal dominance, the characteristic equation analysis, and the alternative definition (criterion) of D–stability turn out to be fruitful, each to a different extent. (The condition based on the Lyapunov theorem looks very theoretical and economically intractable here.) Using the negative diagonal dominance and the alternative definition of D–stability, I have derived the "aggregate economy stability" and the "same sign" sufficient conditions. As for the characteristic equation analysis, I have been able to derive a block of necessary conditions using the negativity of eigenvalues requirement, bypassing the Routh–Hurwitz conditions since they are quite complicated and do not have economic interpretation.

I have studied each group of results in application to the particular setup of models I am working with in order to make the procedure of testing for stability more tractable and at the same time to attach some economic interpretation to this very procedure. The conditions derived are then adapted by me to more simple cases of the general framework considered, namely, univariate economy and structurally homogeneous economy case.

Among the different sufficient conditions and necessary conditions that I have derived, I would like to highlight an easily interpretable unifying condition which is sufficient for convergence of a structurally heterogeneous economy under mixed RLS/SG learning with different degrees of inertia towards a rational expectations equilibrium for a broad class of economic models and a criterion for such a convergence in the univariate case. These conditions are formulated using the concept of a subeconomy and a suitably defined aggregate economy.

The rest of the paper is structured as follows. In the next section I present the environment I am working with: a structurally heterogeneous economy under mixed RLS/SG learning of agents and introduce and explain the concept of δ–stability used to explain the stability of a structurally heterogeneous economy under mixed RLS/SG learning for any (possibly different) degrees of inertia of agents. Section 3 is devoted to sufficient conditions for such a stability, among which are the "aggregate economy" and

[6]See Theorem B.5 in Appendix B.4.

[7]See Theorem B.6 in Appendix B.5.

the "same sign" conditions. In Section 4, I present the necessary conditions for δ–stability[8] that are based on the characteristic equation approach. In Section 5, I demonstrate and provide an interpretation of the derived conditions and the criterion on univariate and multivariate examples, including two specifications of the overlapping generations model and the model of simultaneous markets with structural heterogeneity. Section 6 concludes the paper.

4.2 The Model and Concept of δ–stability

Deriving conditions for stability of a structurally heterogeneous economy under mixed RLS/SG learning for any (possibly different) degrees of inertia of agents, I naturally employ the general framework and notation from Honkapohja and Mitra [43], who were the first to formulate a general criterion for stability of a structurally heterogeneous economy under mixed RLS/SG heterogeneous learning. Structural heterogeneity here means that expectations and learning rules of different agents are different, as well as may be different their fundamental characteristics, such as preferences, endowments, and technology (as opposed to structural homogeneity, which corresponds to the assumption of a representative agent).

"Mixed RLS/SG learning" refers to persistently heterogeneous learning, defined by Honkapohja and Mitra [43] as the one arising when different agents use different types of learning algorithms. In the setup of Honkapohja and Mitra [43] these are RLS and SG algorithms.[9]

The class of linear structurally heterogeneous models with S types of agents with different forecasts is presented by

$$y_t = \alpha + \sum_{h=1}^{S} A_i \hat{E}_t^h y_{t+1} + B w_t, \tag{4.1}$$

$$w_t = F w_{t-1} + v_t, \tag{4.2}$$

where y_t is an $n \times 1$ vector of endogenous variables, w_t is a $k \times 1$ vector of exogenous variables, v_t is a vector of white noise shocks, $\hat{E}_t^h y_{t+1}$ are (in general, non-rational) expectations of the vector of endogenous variables by agent h. It is further assumed that F

[8] The formal definition of this concept is given in the corresponding part of the paper.

[9] More on this (as well as some useful reference for a more detailed study of the terms) can be found in Honkapohja and Mitra [43]. In order not to repeat Honkapohja and Mitra [43], I just briefly present the general setup and the general criterion of stability results. For the full presentation of the RLS/SG learning and the setup, please see Honkapohja and Mitra [43].

($k \times k$ matrix) is such that w_t follows stationary VAR(1) process with $M_w = \lim_{t\to\infty} w_t w_t'$ being a positive definite matrix.

The vector form presented above is a reduced form of the model describing the whole economy, i.e., it is an equation corresponding to the intertemporal equilibrium of the dynamic model. In this model expectations of different agent types influence the current values of endogenous variables.

I also stress the diagonal structure of matrices which I analyze, namely

$$F = diag(\rho_1, ..., \rho_k), M_w = \lim_{t\to\infty} w_t w_t' = diag\left(\frac{\sigma_1^2}{1-\rho_1^2}, ..., \frac{\sigma_k^2}{1-\rho_k^2}\right). \tag{4.3}$$

Structural heterogeneity in the setup of Honkapohja and Mitra [43] is expressed through matrices A_h, which are assumed to incorporate the mass ζ_h of each agent type. That is, $A_h = \zeta_h \cdot \hat{A}_h$, where $\hat{A}_h$ is defined as describing how agents of type h respond to their forecasts. So these are the structural parameters characterizing a given economy. Those may be basic characteristics of agents, like those describing their preferences, endowments, and technology. Structural heterogeneity means that all $\hat{A}_h$'s are different for different types of agents. When $\hat{A}_h = A$ for all h and $\sum \zeta_h = 1$, the economy is structurally homogenous.

In forming their expectations about the next period endogenous variables, agents are assumed to believe that the economic system develops according to the following model, which is called agents' perceived law of motion (PLM).

$$y_t = a_{h,t} + b_{h,t} w_t. \tag{4.4}$$

Mixed learning of agents is introduced as follows. Part of the agents, $h = \overline{1, S_0}$, are assumed to use the RLS learning algorithm, while others, $h = \overline{S_0 + 1, S}$, are assumed to use the SG learning algorithm. Moreover, all of them are assumed to use possibly different degrees of responsiveness to the updating function. These degrees of responsiveness are presented by different degrees of inertia $\delta_h > 0$, which, in formulation of Giannitsarou [37], are constant coefficients before the deterministic decreasing gain sequence in the learning algorithm, which is common for all agents.[10]

After denoting $z_t = (1, w_t)$ and $\Phi_{h,t} = (a_{h,t}, b_{h,t})$, the formal presentation of the learning algorithms in this model can be written as follows.

[10] Honkapohja and Mitra [43] use a more general formulation of degrees of inertia as constant limits in time of the expected ratios of agents' random gain sequences and the common deterministic decreasing gain sequence satisfying certain regularity conditions.

RLS: for $h = \overline{1, S_0}$

$$\Phi_{h,t+1} = \Phi_{h,t} + \alpha_{h,t+1} R_{h,t}^{-1} z_t \left(y_t - \Phi_{h,t}' z_t\right)' \tag{4.5a}$$

$$R_{h,t+1} = R_{h,t} + \alpha_{h,t+1} \left(z_{t-1} z_{t-1}' - R_{h,t}\right) \tag{4.5b}$$

SG: for $h = \overline{S_0 + 1, S}$

$$\Phi_{h,t+1} = \Phi_{h,t} + \alpha_{h,t+1} z_t \left(y_t - \Phi_{h,t}' z_t\right)'. \tag{4.6}$$

Honkapohja and Mitra [43] show that stability of the REE, Φ_t, in this model is determined by the stability of the ODE:

$$\frac{d\Phi_h}{d\tau} = \delta_h \left(T(\Phi')' - \Phi_h\right), h = \overline{1, S_0} \tag{4.7}$$

$$\frac{d\Phi_h}{d\tau} = \delta_h M_z \left(T(\Phi')' - \Phi_h\right), h = \overline{S_0 + 1, S}, \tag{4.8}$$

where $M_z = \lim_{t\to\infty} E z_t z_t' = \begin{pmatrix} 1 & 0 \\ 0 & M_w \end{pmatrix}$ and $T(\Phi')$ is a mapping of the PLM parameters into the parameters of the actual law of motion (ALM)

$$y_t = \left[\alpha + \sum_{h=1}^{S} A_h a_{h,t}, \left(\sum_{h=1}^{S} A_h b_{h,t}\right) F + B\right] \begin{bmatrix} 1 \\ w_t \end{bmatrix} = T(\Phi') z_t,$$

which is obtained when one plugs the forecast functions based on the agents' PLMs (4.4)

$$\hat{E}_t^h y_{t+1} = a_{h,t} + b_{h,t} F w_t \tag{4.9}$$

into the reduced form of the model (4.1) and (4.2). So,[11]

$$T(a_{h,t}, b_{h,t}) = \left(\alpha + \sum_{h=1}^{S} A_h a_{h,t}, \left(\sum_{h=1}^{S} A_h b_{h,t}\right) F + B\right). \tag{4.10}$$

The conditions for stability of this ODE give the general criterion of stability for this class of models presented in Proposition 5 in Honkapohja and Mitra [43], introduced (without proof) here for the reader's convenience.

Criterion 4.1 *(Proposition 5 in Honkapohja and Mitra [43]) In the economy (4.1) and (4.2), mixed RLS/SG learning converges globally (almost surely) to the minimal state*

[11] For details, please see Honkapohja and Mitra [43].

variable (MSV)[12] *solution if and only if matrices* $D_1\Omega$ *and* $D_w\Omega_F$ *have eigenvalues with negative real parts, where*

$$D_1 = \begin{pmatrix} \delta_1 I_n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \delta_S I_n \end{pmatrix}, \Omega = \begin{pmatrix} A_1 - I_n & \cdots & A_S \\ \vdots & \ddots & \vdots \\ A_1 & \cdots & A_S - I_n \end{pmatrix} \tag{4.11}$$

$$D_w = \begin{pmatrix} D_{w1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{wS} \end{pmatrix}, \quad \begin{matrix} D_{wh} = \delta_h I_{nk}, h = \overline{1, S_0} \\ D_{wh} = \delta_h \left(M_w \otimes I_n\right), h = \overline{S_0 + 1, S} \end{matrix}$$

$$\Omega_F = \begin{pmatrix} F' \otimes A_1 - I_{nk} & \cdots & F' \otimes A_S \\ \vdots & \ddots & \vdots \\ F' \otimes A_1 & \cdots & F' \otimes A_S - I_{nk} \end{pmatrix},$$

with $\otimes$ denoting the Kronecker product.

In the "diagonal" environment I consider, the problem of finding conditions for stability of both $D_1\Omega$ and $D_w\Omega_F$ under any (possibly different) degrees of inertia of agents, $\delta > 0$, is simplified to finding stability conditions of $D_1\Omega$ and $D_1\Omega_{\rho_l}$,where Ω_{ρ_l} is obtained from Ω by substituting all A_h with $\rho_l A_h$, where $|\rho_l| < 1$ as w_t follows stationary VAR(1) process by the setup of the model (see Appendix B.6 for a more detailed proof of Proposition 4.2 below).

$$\Omega_{\rho_l} = \begin{pmatrix} \rho_l A_1 - I_n & \cdots & \rho_l A_S \\ \vdots & \ddots & \vdots \\ \rho_l A_1 & \cdots & \rho_l A_S - I_n \end{pmatrix}, \forall l = 0, ..., k, (\rho_0 = 1). \tag{4.12}$$

Proposition 4.2 *(The criterion for stability of a structurally heterogeneous economy under mixed RLS/SG learning for the diagonal environment case under any (possibly different) degrees of inertia of agents,* $\delta > 0$*). In the structurally heterogeneous economy (4.1), (4.2) and (4.3), mixed RLS/SG learning (4.5), (4.6) and (4.9) converges globally (almost surely) to an MSV REE solution for any (possibly different) degrees of inertia of*

[12] As it is mentioned in ch. 8 of Evans and Hokapohja [29], the concept of the MSV solution was introduced by McCallum [57] for linear rational expectations models. As is defined in Evans and Honkapohja [29], this is the solution that depends linearly on a set of variables (in our case it is the vector of exogenous variables and the intercept); this solution is such that there is no other solution that depends linearly on a smaller set of variables.

agents, $\delta > 0$, if and only if matrices $D_1\Omega_{\rho l}$ are stable for any $\delta > 0$, where D_1 and Ω_{ρ_l} are defined in (4.11) and (4.12), respectively.

Proof. See Appendix B.6. □

I also use the special blocked—diagonal structure of matrix D_1 which is the feature of the dynamic environment in this class of models. In a sense these positive diagonal D–matrices now may be called positive blocked—diagonal δ–matrices. It allows me to formulate the concept of δ–stability by analogy to the terminology of the concept of D–stability, studied for example in Johnson (1974).

Definition 4.1 *Given n, the number of endogenous variables, and S, the number of agent types,* ***δ–stability*** *is defined as stability of the structurally heterogeneous economy (4.1) and (4.2) under mixed RLS/SG learning (4.5), (4.6) and (4.9) under any (possibly different) degrees of inertia of agents, $\delta > 0$.*

δ–stability, thus formulated, has the same meaning in models with heterogeneous learning described above as has the E–stability condition in models with homogeneous RLS learning. The $E-$stability condition is a condition for asymptotic stability of an REE under homogeneous RLS learning. The REE of the model is stable if it is locally asymptotically stable under the following ODE:

$$\frac{d\theta}{d\tau} = T(\theta) - \theta, \tag{4.13}$$

where θ are the estimated parameters from agents PLMs, the T–map is defined in (4.10), and τ is a "notional" ("artificial") time. The fixed point of this ODE is the REE of the model.

4.3 Sufficient Conditions for δ–stability

4.3.1 Aggregate Economy Conditions

Following the description of the approaches to stability in the introduction, I now separately consider each of them. First, I follow the **negative diagonal dominance** approach and it allows me to show that in the setting specified above δ–stability depends on E–stability of the aggregate economy which is the upper boundary of aggregate economies with weights of aggregation across agents, ϕ, and weights of aggregation across endogenous variables, ψ.

I have been encouraged by the result that follows from Propositions 2 and 3 in Honkapohja and Mitra [43] that for stability under heterogeneous RLS or SG learning with the same degrees of inertia, stability in the economy aggregated across agent types (average economy) turns out to be crucial. Following Honkapohja and Mitra [43], who aggregated an economy across agents by introducing the concept of average (aggregated across agents) economy, I also began to look for the concept of an aggregate economy that has to be crucial for stability of a structurally heterogeneous economy under mixed RLS/SG heterogeneous learning with different degrees of inertia of agents. The basic idea is that there has to be a way to aggregate an economy in an economically reasonable way, so that E–stability in the aggregate economy is sufficient for δ–stability in the original economy.

I proceed with aggregation of the economy starting from the following aggregation across agents used by Honkapohja and Mitra [43]:

$$y_t = \alpha + A^M \hat{E}_t^{AV} y_{t+1} + Bw_t.$$

It turns out that it is convenient, in addition to the aggregation across agents above, to consider aggregation across endogenous variables. The economy aggregated across endogenous variables will no longer be a vector but a scalar, which means that it can characterize many economies.

I rewrite the formulas used by Honkapohja and Mitra [43] for average expectations as

$$\begin{aligned} E_t^{AV} y_{t+1} &= (A^M)^{-1} \left(S \sum_{h=1}^{S} \frac{1}{S} A_h E_t^h y_{t+1} \right) \\ A_M &= S \sum_{h=1}^{S} \frac{1}{S} A_h = S \sum_{h=1}^{S} \frac{1}{S} \zeta_h \hat{A}_h. \end{aligned}$$

After this, one can interpret the aggregation done by Honkapohja and Mitra [43] as follows: first, one takes the weight of each agent type in calculating aggregate expectations of one representative agent to be equal to $\frac{1}{S}$ and then multiplies these expectations by S in order to be consistent with the model that consists of S types of agents. (So that the size of the economy is preserved by replacing each type of agent with a representative agent).

In general, when aggregating expectations one may use different weights for different types of agents that sum up to one in order to reflect the relative importance of a particular agent type expectation in the aggregate economy. So, in my aggregation, I first create a representative agent type by averaging across all agent types (assigning a weight to each type and summing over all types) and then I aggregate over all types by multiplying the representative (average) agent type by S in order to preserve the size of the aggregate economy.[13]

If I write the aggregate economy using different weights for aggregation of expectations across agents, I will get

$$\hat{E}_t^{Weighted} y_{t+1} = \left(A^{Weigted}\right)^{-1} \left(S \sum_{h=1}^{S} \phi_h A_h \hat{E}_t^h y_{t+1}\right)$$
$$A^{Weigted} = S \sum_{h=1}^{S} \phi_h A_h = S \sum_{h=1}^{S} \phi_h \zeta_h \hat{A}_h,$$

where $\phi_h > 0$, $h = \overline{1, S}$ are weights of single agent types used in calculating aggregate expectations, such that $\sum_{h=1}^{S} \phi_h = 1$.

Next, given the weights of aggregation across endogenous variables $\psi_i > 0$, $\sum_{i=1}^{n} \psi_i = 1$, and across agent types $\phi_h > 0$, $\sum_{h=1}^{S} \phi_h = 1$ (and denoting a_{ij}^h the element in the i^{th} row and j^{th} column of matrix A_h), I aggregate the economy in the following way

$$y_t^{AG} = \sum_i \psi_i y_{it} = \sum_i \psi_i \alpha_i + \sum_h S\phi_h \sum_i \psi_i \sum_j a_{ij}^h \hat{E}_t^h y_{jt+1} + \left(\sum_i \psi_i B^i\right) w_t =$$
$$= \sum_i \psi_i \alpha_i + \beta^{AG}(\psi, \phi) \hat{E}_t^{AG}\left(y_{t+1}^{AG}\right) + \left(\sum_i \psi_i B^i\right) w_t, \text{ where}$$

$$\beta^{AG}(\psi, \phi) = S \sum_h \phi_h \sum_i \psi_i \sum_j a_{ij}^h, \tag{4.14}$$

$$\hat{E}_t^{AG}\left(y_{t+1}^{AG}\right) = \Big(\sum_{h=1}^{S} S\phi_h \underbrace{\sum_i \psi_i \sum_j a_{ij}^h}_{\beta_h}\Big)^{-1} \sum_{h=1}^{S} S\phi_h \underbrace{\sum_i \psi_i \sum_j a_{ij}^h}_{\beta_h} \hat{E}_t^h y_{jt+1}, \tag{4.15}$$

and B^i denotes the i^{th} row of B. So, using the derivations above I formulate the following definition.

[13] The new dimension in weighting agent types, in addition to the mass of each agent type ζ_h incorporated in matrices A_h, may also have the following interpretation. I can assume that the share of each agent type expectation in the average expectations of the population is determined not only by their mass in the population (their physical share), but also by each type's influence, other than their share in the population (e.g., political or mass media power or other type of influence in the social life of the whole population). By assigning additional weights to each agent type I provide a measure of the share of influence of each agent type in the overall expectations, bypassing the intermediate step of measuring the influence of each agent type on other agent types separately.

Definition 4.2 *Given the weights of aggregation across endogenous variables* $\psi_i > 0$, $\sum_{i=1}^{n} \psi_i = 1$, *and across agent types* $\phi_h > 0$, $\sum_{h=1}^{S} \phi_h = 1$, ***the aggregate economy*** *for an economy described by (4.1), (4.2) and (4.3) is defined as*

$$\begin{aligned} y_t^{AG} &= \sum_i \psi_i \alpha_i + \beta^{AG}(\psi, \phi) \hat{E}_t^{AG}\left(y_{t+1}^{AG}\right) + \left(\sum_i \psi_i B^i\right) w_t, \\ & \quad (4.2) \text{ and } (4.3), \end{aligned}$$

where $\beta^{AG}(\psi, \phi)$ *and* $E^{AG}\left(y_{t+1}^{AG}\right)$ *are defined in (4.14) and (4.15) respectively.*

It is also useful to consider an economy that bounds above all possible economies with all possible combinations of signs of a_{ij}^h aggregated using weights ψ and ϕ. This is obviously our original aggregate model written in absolute values. When all elements in the model, a_{ij}^h, endogenous variables and their expectations are positive, this limiting model exactly coincides with the model considered. So, this is an attainable supremum. Thus I have the following limiting aggregate model:

$$y_t^{AG} = \sum_i \psi_i y_{it} \le y_t^{AG\,\mathrm{mod}} = \sum_i \psi_i \left|y_{it}\right| \le$$

$$\le \sum_i \psi_i \left|\alpha_i\right| + \beta^{AG\,\mathrm{mod}}(\psi, \phi) \hat{E}_t^{AG\,\mathrm{mod}}\left(y_{t+1}^{AG\,\mathrm{mod}}\right) + \left|\left(\sum_i \psi_i B^i\right) w_t\right|, \text{ where}$$

$$\beta^{AG\,\mathrm{mod}}(\psi, \phi) = S \sum_h \phi_h \sum_i \psi_i \sum_j \left|a_{ij}^h\right| \tag{4.16}$$

$$\hat{E}_t^{AG\,\mathrm{mod}}\left(y_{t+1}^{AG\,\mathrm{mod}}\right) = \Big(\sum_{h=1}^{S} S\phi_h \underbrace{\sum_i \psi_i \sum_j \left|a_{ij}^h\right|}_{\beta_h}\Big)^{-1} \sum_{h=1}^{S} S\phi_h \underbrace{\sum_i \psi_i \sum_j \left|a_{ij}^h\right|}_{\beta_h} \hat{E}_t^h \left|y_{jt+1}\right| \tag{4.17}$$

Definition 4.3 *Given the weights of aggregation across endogenous variables* $\psi_i > 0$, $\sum_{i=1}^{n} \psi_i = 1$, *and across agent types* $\phi_h > 0$, $\sum_{h=1}^{S} \phi_h = 1$, ***the limiting aggregate economy*** *for an economy described by (4.1), (4.2) and (4.3) is defined as*

$$\begin{aligned} y_t^{AG\,\mathrm{mod}} &= \sum_i \psi_i \left|\alpha_i\right| + \beta^{AG\,\mathrm{mod}}(\psi, \phi) \hat{E}_t^{AG\,\mathrm{mod}}\left(y_{t+1}^{AG\,\mathrm{mod}}\right) + \left|\left(\sum_i \psi_i B^i\right) w_t\right|, \\ & \quad (4.2) \text{ and } (4.3), \end{aligned}$$

where $\beta_{abs}^{AG}(\psi, \phi)$ *and* $E_{abs}^{AG}\left(y_{abst+1}^{AG}\right)$ *are defined in (4.16) and (4.17) respectively.*

Remark 4.1 *If this limiting aggregate economy is E–stable, then all corresponding aggregate economies with various combinations of signs of* a_{ij}^h *are E–stable.*

The structure of this limiting aggregate coefficient $\beta^{AG\,\mathrm{mod}}$ is as follows. $\sum_i \psi_i \left|a_{ij}^h\right|$ is the coefficient before the expectation of endogenous variable j in the aggregate economy composed of one single agent type h. Notice that this coefficient is calculated for the expectation of endogenous variable j, that enters the aggregate product with coefficient ψ_j. So, I may name the ratio $\sum_i \psi_i \left|a_{ij}^h\right| / \psi_j$ **endogenous variable j "own" expectations relative coefficient.** By looking at the values of these coefficients I will be able to judge the weight a particular agent type has in the economy in terms of the aggregate β–coefficient. The next proposition is formulated in terms of these relative coefficients and stresses the fact that weights of agents in calculating aggregate expectations have to be put into accordance with this economic intuition in order to have stability under heterogeneous learning.

Proposition 4.3 *If there exists at least one pair of vectors of weights for aggregation of endogenous variables ψ and weights ϕ for aggregation of agents such that for each agent every endogenous variable's "own" expectations relative coefficient is less than the weight of the agent used in calculating aggregate expectations, i.e.* $\sum_i \psi_i \left|a_{ij}^h\right| / \psi_j < \phi_h, \forall j, \forall h,$ *then the economy described by (4.1), (4.2) and (4.3) is δ–stable.*

Proof. See Appendix B.6. □

But this proposition above does not give a real rule of thumb (as it implies looking for systems of weights) that could be used to say if a particular economy is stable under heterogeneous learning. For this purpose I have constructed four maximal aggregate β–coefficients described below. If they are less than one, the economic system is δ–stable.

Thus I go even further looking for upper boundaries by considering not only any possible signs of a_{ij}, but also values of weights ψ and ϕ. These boundaries can be derived for four different subsets of aggregate economies depending on the values of weights ψ and ϕ: with arbitrary weights of agents and endogenous variables, and with either equal weights of agents $\frac{1}{S}$ or equal weights of endogenous variables $\frac{1}{n}$, or both. So, each aggregate economy from a particular subset of aggregate economies is bounded above by the following **maximal aggregate economy**

$$y_t^{AG} = \sum_i \psi_i y_{it} \le y_t^{AG\,\mathrm{mod}} = \sum_i \psi_i \left|y_{it}\right| \le y_t^{AG\,\max} =$$

$$= \sum_i \psi_i \left|\alpha_i\right| + \beta_r^{AG\,\max} \hat{E}_t^{AG\,\max}\left(y_{t+1}^{AG\,\max}\right) + \left|\left(\sum_i \psi_i B^i\right) w_t\right|,$$ where $\beta_r^{AG\,\max}$ is defined in Table 4.1.

	$r=1$	$r=2$	$r=3$	$r=4$
Subset	ψ–any, ϕ–any	ψ–any, $\phi=\frac{1}{S}$	$\psi=\frac{1}{n}$, ϕ–any	$\psi=\frac{1}{n}$, $\phi=\frac{1}{S}$
$\beta_r^{AG\max}=$	$S\sum\limits_j \max\limits_{h,i}\left\|a_{ij}^h\right\|$	$\max\limits_i \sum\limits_h \sum\limits_j \left\|a_{ij}^h\right\|$	$S\sum\limits_i \max\limits_{h,j}\left\|a_{ij}^h\right\|$	$\sum\limits_h \max\limits_j \sum\limits_i \left\|a_{ij}^h\right\|$

Table 4.1: Maximal aggregate β–coefficients.

First, let us prove that these **maximal aggregate β–coefficients** are actually upper boundaries for $\beta^{AG\,\mathrm{mod}}(\psi,\phi) = S\sum\limits_h \phi_h \sum\limits_i \psi_i \sum\limits_j \left|a_{ij}^h\right|$ for different subsets of aggregate economies. Formally, the result can be written in a form of the following proposition.

Proposition 4.4 *Maximal aggregate β–coefficients defined in Table 4.1 are upper boundaries for $\beta^{AG\,\mathrm{mod}}(\psi,\phi) = S\sum\limits_h \phi_h \sum\limits_i \psi_i \sum\limits_j \left|a_{ij}^h\right|$ for the corresponding subsets of aggregate economies.*

Proof. See Appendix B.6. □

Now, it is possible to give the formal definition of **the maximal aggregate economy**.

Definition 4.4 *Given the weights of aggregation across endogenous variables $\psi_i > 0$, $\sum\limits_{i=1}^{n}\psi_i = 1$, and across agent types $\phi_h > 0$, $\sum\limits_{h=1}^{S}\phi_h = 1$, **the maximal aggregate economy** for an economy described by (4.1), (4.2) and (4.3) is defined as*

$$\begin{aligned} y_t^{AG\max} &= \sum_i \psi_i \left|\alpha_i\right| + \beta_r^{AG\max}(\psi,\phi)\,\hat{E}_t^{AG\max}\left(y_{t+1}^{AG\max}\right) + \left|\left(\sum_i \psi_i B^i\right) w_t\right|, \\ & \quad (4.2) \text{ and } (4.3), \end{aligned}$$

where $\beta_r^{AG\max}(\psi,\phi)$ is defined Table 4.1 and $\hat{E}_t^{AG\max}\left(y_{t+1}^{AG\max}\right)$ is defined to be equal to $\hat{E}_t^{AG\,\mathrm{mod}}\left(y_{t+1}^{AG\,\mathrm{mod}}\right)$ in (4.17).

Notice that each of the above–described boundaries is constructed in such a way that it does not replicate the boundary for a broader set of aggregate models to which this particular model belongs. It is possible to do so by applying the max operator to different groupings of elements of sum and it becomes possible only for a particular subset of aggregate models and which was not possible to apply to a broader set. Under equal

$\left|a_{ij}^{h}\right| = |a|$ all these maximal aggregate β–coefficients coincide with $\beta^{AG\,\mathrm{mod}}(\psi, \phi) = nS\,|a|$. So, these are attainable maxima.

Thus I have managed to aggregate the economy into one dimension and to find the maximal aggregate economies that bound all of such aggregate economies within a particular subset. If one of these maximal aggregate economies is E–stable (i.e. if at least one of the maximal aggregate β–coefficients is less than one), then all aggregate subeconomies from a particular subset of aggregate economies are E–stable. As I have already mentioned the concept of a subeconomy, I shall now introduce its formal definition as this concept is convenient to use in proofs and conditions for δ–stability.

Definition 4.5 *A subeconomy $(h_1, ..., h_p)$ of size p for an economy (4.1) and (4.2) is defined as consisting only of a part of agents from the original economy:*

$$y_t = \alpha + \sum_{k=1}^{p} A_{i_k} \hat{E}_t^h y_{t+1} + B w_t, \tag{4.18}$$

$$w_t = F w_{t-1} + v_t, \tag{4.19}$$

where $(h_1, ..., h_p) \subseteq (1, ..., S)$ is a set of numbers of agent types present in the subeconomy. A single economy is a particular case of a subeconomy with only one type of agent.

Now I am ready to formulate the result which stresses the key role of E–stability in the aggregate economy in stability of the original structurally heterogeneous economy under mixed RLS/SG learning with possibly different degrees of inertia (recall Proposition 2 and Proposition 3 in Honkapohja and Mitra [43]). The key result is as follows.

Proposition 4.5 *If one of the maximal aggregate economies is E–stable (i.e., one of the maximal aggregate β–coefficients is less than one), then the economy (4.1), (4.2) and (4.3) is $\delta-$stable. Notice that all subeconomies are also δ–stable under this condition.*

Proof. See Appendix B.6. □

This result gives a direct rule how to construct δ–stable economies. I think that this is quite a strong result that says that there is one economic unifying condition (such as aggregate β–coefficient less than one) such that when it holds true all the economies with the same absolute values of a_{ij}^{h} (with all possible combinations of their signs) are δ–stable.

This condition shows how robust is the stability of a model to a change in sign of some coefficients in the economy during the time. Also, fixing certain components in these aggregate β–coefficients, I may see how the value of other coefficients is flexible for the economy to remain δ–stable. This can be useful in the case when one does not know the exact sign of some coefficient in matrix A_h, but may estimate that its absolute value belongs to some interval with some probability (the situation typical for statistical interval estimation). As an example, the policy maker may know some structural coefficients in the economy and have to choose some parameters itself (like the ones for the policy rule). This formula allows it to see what is the range of parameters it may choose in order to make sure that the economy is δ–stable.

It is possible to simplify the derived conditions for more simple cases, namely, for a univariate model and a structurally homogeneous model.

Proposition 4.6 *A univariate ($n = 1$) economy described by (4.1) and (4.2) is δ–stable for any combination of signs of coefficients if and only if* $|A_1| + |A_2| + ... + |A_s| < 1$.

Proof. Obvious: the necessary condition for Ω to be stable under any δ is $A_1 + ... + As < 1$ It follows from the condition on the determinant of $-\Omega$, which has to be positive. This determinant equals $-(A_1 + ... + As) + 1$. For the above condition to hold true for any signs of $A_h, h = \overline{1, S}$, it is necessary and sufficient that $|A_1| + |A_2| + ... + |A_s| < 1$. □

Proposition 4.7 *For a structurally homogeneous economy:* $A_h = \zeta_h A$, $\zeta_h > 0, \sum_{h=1}^{S} \zeta_h = 1$; *to be δ–stable it is sufficient that at least one of the following maximal aggregate β–coefficients be less than one;* $\max_i \sum_j |a_{ij}|$ *and* $\max_j \sum_i |a_{ij}|$.

Proof. Direct application of Proposition 4.5. □

4.3.2 "Same Sign" Conditions

Following the steps of the proof of observation (iv) in Johnson [44] (the formulation of this observation is presented in Appendix B.5), which is, in fact, the alternative definition of D–stability, I get an alternative definition of blocked—diagonal (D_b)–stability, that is stability of $D_b\Omega$ for any positive blocked—diagonal matrix D_b. This alternative definition of D_b–stability is then used to derive conditions for δ–stability.

Definition 4.6 *(D_b–stability) Matrix A of size $nS \times nS$ is D_b-stable if D_bA is stable for any positive blocked–diagonal matrix $D_b = diag(\delta_1, ..., \delta_1, ..., \delta_S, ..., \delta_S)$.*

Proposition 4.8 *(Alternative definition of D_b–stability). Consider $M_{nS}(C)$, the set of all complex $nS \times nS$ matrices, D_{bnS}, the set of all $nS \times nS$ blocked–diagonal matrices with positive diagonal entries. Take $A \in M_{nS}(C)$ and suppose that there is $F \in D_{bnS}$ such that FA is stable. Then A is D_b–stable if and only if $A \pm iD_b$ is non–singular for all $D_b \in D_{bnS}$. If $A \in M_{nS}(R)$, – the set of all $nS \times nS$ real matrices, then "$\pm$" in the above condition may be replaced with "$+$" since, for a real matrix, any complex eigenvalues come in conjugate pairs.*

Proof. (The proof is just a modification of the proof of observation (iv) in Johnson [44] for my blocked–diagonal case) **Necessity.** Let A be D_b–stable, that is EA is stable for all positive blocked–diagonal $E \in D_{bnS}$. This means that $\pm i$ cannot be an eigenvalue of matrix EA for any $E \in D_{bnS}$. That is $EA \pm iI$ is non–singular for all $E \in D_{bnS}$, or $A \pm iD_b$ is non–singular for all $D_b = E^{-1} \in D_{bnS}$. **Sufficiency**. By contradiction. Let A be not D_b–stable. Thus, I have that there exists some $E \in D_{bnS}$ such that FA is stable, while EFA is not stable. By continuity, it follows that either value, $\pm i$, is an eigenvalue of $\frac{1}{\alpha}(tE + (1-t)I) FA$ for some $0 < t \leq 1$ and $\alpha > 0$. So, $A \pm iD_b$ is singular for $D_b = \alpha F^{-1} (tE + (1-t) I)^{-1} \in D_{bnS}$. Contradiction. □

Taking F as an identity matrix, and D as $diag(\frac{1}{\delta_1}, ..., \frac{1}{\delta_1}, ..., \frac{1}{\delta_S}, ..., \frac{1}{\delta_S})$, $\delta_h > 0, h = \overline{1, S}$, in the above proposition, I get the following necessary and sufficient condition (criterion) for δ–stability:

Proposition 4.9 *(criterion for δ–stability in terms of structural and learning heterogeneity) An economy described by (4.1), (4.2) and (4.3) is δ–stable if and only if the corresponding matrix Ω, defined in (4.11), is stable and*

$$\det\left[\sum_{h=1}^{S}\left(\frac{-\rho_l A_h}{1+\frac{i}{\delta_h}}\right) + I\right] = \det\left[\left(\sum_{h=1}^{S}\frac{1}{1+\frac{1}{\delta_h^2}}(-\rho_l A_h) + I\right) \pm i\left(\sum_{h=1}^{S}\frac{\frac{1}{\delta_h}}{1+\frac{1}{\delta_h^2}}(-\rho_l A_h)\right)\right] \neq 0$$
$$\forall \delta_h > 0, h = \overline{1, S}, \forall l = 0, 1, ..., k, (\rho_0 = 1)$$

For the univariate case ($n = 1$) this condition simplifies to Ω — stable and

$$\left(\sum_{h=1}^{S}\frac{1}{1+\frac{1}{\delta_h^2}}(-\rho_l A_h) + 1\right) \neq 0 \text{ or } \sum_{h=1}^{S}\frac{\frac{1}{\delta_h}}{1+\frac{1}{\delta_h^2}}(-\rho_l A_h) \neq 0,$$
$$\text{or both}, \forall \delta_h > 0, h = \overline{1, S}, \forall l = 0, 1, ..., k, (\rho_0 = 1).$$

The **alternative definition of D–stability** approach allows us to derive the "same sign" conditions for the cases $n = 1, 2$ and necessary and sufficient conditions for δ–stability for $n = 1$.

Proposition 4.10 *(Criterion for δ–stability in the univariate case in terms of structural heterogeneity only) In the case $n = 1$, an economy described by (4.1), (4.2) and (4.3) is δ–stable if and only if the corresponding matrix Ω, defined in (4.11), is stable and at least one of the following holds true: the same sign condition (all A_h are greater than or equal to zero and at least one is strictly greater than zero or all A_h are less than or equal to zero and at least one is strictly less than zero), or all average economies with $A_{(h_1,...,h_p)} = \sum\limits_{(h_1,...,h_p)} A_h$.corresponding to subeconomies $(h_1, ..., h_p)$ of all sizes p are not E–unstable and for each $l = 0, 1, ..., k$ $(\rho_0 = 1)$ there exists at least one average economy corresponding to subeconomy $(h_1^*(l), ..., h_p^*(l))$ in each size p for which the stability coefficient $\sum\limits_{(h_1^*(l),...,h_p^*(l))} \rho_l A_i$ is strictly less than one.*

Remark 4.2 *Due to Proposition 2 of Honkapohja and Mitra [43], E–stability/instability of a particular average economy is necessary and sufficient for stability/instability of the corresponding subeconomy under transiently heterogeneous SG learning, which is determined by the stability of matrix $\Omega_{(h_1,..,h_p)}$. So, using this criterion, one may use interchangeably the conditions for the stability of average economies or the conditions for stability of subeconomies, whatever is more convenient in a particular setting. By the same Proposition 2 of Honkapohja and Mitra [43], the condition for stability of matrix Ω can also be considered as a condition for E–stability of the "largest" (including all agents in calculating the average coefficient) average economy corresponding to the original economy.*

Proposition 4.11 *In the case $n = 2$, the economy described by (4.1), (4.2) and (4.3) is δ–stable if the corresponding matrix Ω, defined in (4.11), is stable and the following "same sign" condition holds true:*

$$\det(-\rho_l A_i) \geq 0, [\det mix\,(-\rho_l A_i, -\rho_l A_j) + \det mix\,(-\rho_l A_j, -\rho_l A_i)] \geq 0, i \neq j, M_1(-\rho_l A_i) \geq 0$$

or

$$\det(-\rho_l A_i) \leq 0, [\det mix\,(-\rho_l A_i, -\rho_l A_j) + \det mix\,(-\rho_l A_j, -\rho_l A_i)] \leq 0, i \neq j, M_1(-\rho_l A_i) \leq 0,$$

$$\forall l = 0, 1, ..., k, (\rho_0 = 1),$$

where $mix\,(-\rho_l A_i, -\rho_l A_j)$ *denotes a matrix of structural parameters of a pairwise–mixed economy and is composed by mixing columns of a pair of matrices* $\rho_l A_i$, $\rho_l A_j$, *for any* i, $j = \overline{1, S}$.

Proof. See Appendix B.6. □

Remark 4.3 *Unfortunately, though similar "same sign" conditions naturally follow from the alternative definition of* D*–stability for cases* $n > 2$, *stability of* Ω *and a similar "same sign" condition are not sufficient for* δ*–stability in this case. For example, a similar "same sign" condition for case* $n = 3$ *looks like*

$$\begin{aligned} M_3\,(mix\,(-\rho_l A_i, -\rho_l A_j, -\rho_l A_k)) &> 0, M_2(mix\,(-\rho_l A_i, -\rho_l A_j)) > 0, M_1(-\rho_l A_i) > 0 \\ & \quad or \\ M_3\,(mix\,(-\rho_l A_i, -\rho_l A_j, -\rho_l A_k)) &< 0, M_2(mix\,(-\rho_l A_i, -\rho_l A_j)) < 0, M_1(-\rho_l A_i) < 0, \\ \forall l &= 0, 1, ..., k(\rho_0 = 1) \end{aligned}$$

Here, the $M_n(mix())$ *operator means the sum of all possible principal minors of size* n *of a particular mix between matrices.*

4.4 Necessary Conditions for δ–stability

The **characteristic equation** approach (which in my formulation leaves aside the intractable Routh–Hurwitz conditions) has allowed me to derive strong necessary conditions for δ–stability that provide an easy test for non–δ–stability of the model. Note that necessary conditions do not require a diagonal structure of F and M_w.

Condition (⋆) *All sums of the same-size principal minors of* $D_{1r}\,(-\Omega_r)$ *are nonnegative for all subeconomies* $r = (h_1, ..., h_p)$ *for all* p *for all positive block–diagonal matrices* D_{1r}, *where* D_{1r} *and* Ω_r *defined similar to* D_1 *and* Ω *in (4.11) correspond to a subeconomy of the economy under consideration.*

Proposition 4.12 *Necessary condition for* δ*–stability: For the economy (4.1) and (4.2) to be* δ*–stable, it is necessary that Condition (⋆) holds true.*

Proof. See Appendix B.6. □

The condition above can not be used as a test for non–δ–stability, as it requires checking all subeconomies' sums of minors for all possible D_{1r}. That is why below I have constructed a condition that has a direct testing application.

Proposition 4.13 *Necessary condition for δ–stability: For the economy (4.1) and (4.2) to be δ–stable, it is necessary that all sums of the same-size principal minors of minus matrices corresponding to subeconomies $(-\Omega_r)$ be non-negative for each corresponding subeconomy $r = (h_1, ..., h_p)$.*

Proof. See Appendix B.6. □

I think that this is quite a strong necessary condition, which implies that a lot of models will not satisfy it, and will not be δ–stable. Note that stability of each single economy and subeconomies is a sufficient condition for the condition above to hold true. A weaker requirement that all subeconomies be not unstable (non-positive real parts of eigenvalues) is also sufficient.

4.5 Economic Examples

4.5.1 Univariate Case

I exploit the same reduced form used as an example of a univariate model in Honkapohja and Mitra [43]. Such a reduced form can be a result of equilibrium in a non-stochastic basic overlapping generations model (so-called Samuelson model) developed in Chapter 4 of Evans and Honkapohja [29]. Here I develop it for the heterogeneous agents case.

There are S types of agents in the economy, each of whom lives for two periods (young and old). Population is constant: old agents who died in the second period are replaced with the same number of young agents in the next period. When agents are young they work supplying labor $n_{h,t}$ and save the revenue obtained from working; when they are old, they consume their savings in amount $c_{h,t+1}$. Output equals labor supply, so wage earned equals the same period price of the consumption good. There is a constant stock of money, M , which is the only means of saving in the economy. So, in a non-autarky case, there is trade in the economy between generations: each period t, output produced by the young generation is sold to the old agents on a competitive market using money.

Each agent h born at time t has a constant elasticity of substitution utility function

$$U_h(c_{h,t+1}, n_{h,t}) = \frac{(c_{h,t+1})^{1-\sigma_h}}{1-\sigma_h} - \frac{(n_{h,t})^{1+\varepsilon_h}}{1+\varepsilon_h}, \sigma_h, \varepsilon_h > 0.$$

Budget constraints for the first and second periods of the life of agent h are

$$\begin{aligned} p_t n_{h,t} &= M_{h,t} \\ p^{h,e}_{t+1} c_{h,t+1} &= M_{h,t}, \end{aligned}$$

respectively, where p_t is the price of the good and $M_{h,t}$ denotes the nominal savings of agent h after the first period. $p^{h,e}_{t+1}$ denotes expectations of the next period price made today (they are taken to be point expectations, as the economy considered is non–stochastic).

After solving the agent's problem, the (real) saving function of the agent looks like

$$F_h\left(\frac{p^{h,e}_{t+1}}{p_t}\right) \equiv \left(\frac{p^{h,e}_{t+1}}{p_t}\right)^{\frac{\sigma-1}{\sigma+\varepsilon}} = \frac{M_{h,t}}{p_t}.$$

The market clearing condition equates total savings to the stock of money in the economy each period

$$\frac{M}{p_t} = \sum_{h=1}^{S} F_h\left(\frac{p^{h,e}_{t+1}}{p_t}\right) = \sum_{h=1}^{S}\left(\frac{p^{h,e}_{t+1}}{p_t}\right)^{\frac{\sigma-1}{\sigma+\varepsilon}}, \text{ or}$$

$$H(p_t, (p^{h,e}_{t+1})^S_{h=1}) \equiv \frac{M}{p_t} - \sum_{h=1}^{S} F_h\left(\frac{p^{h,e}_{t+1}}{p_t}\right) = \frac{M}{p_t} - \sum_{h=1}^{S}\left(\frac{p^{h,e}_{t+1}}{p_t}\right)^{\frac{\sigma-1}{\sigma+\varepsilon}} = 0. \tag{4.20}$$

I use Taylor expansion to linearize this condition around the steady state $p_{t+1} = p_t = \bar{p} = M/S$,

$$\tilde{p}_t = \sum_{h=1}^{S} \underbrace{\left[-\frac{\partial H}{\partial p_t}\frac{\partial H}{\partial p^{h,e}_{t+1}}\bigg|_{p_{t+1}=p_t=\bar{p}} \right]}_{A_h} \tilde{p}^{h,e}_{t+1}.$$

$\frac{\partial H}{\partial p_t}$ and $\frac{\partial H}{\partial p^{h,e}_{t+1}}$ could be easily calculated using (4.20) and evaluated at the steady state. Thus,

$$\begin{aligned} \frac{\partial H}{\partial p_t}\bigg|_{p_{t+1}=p_t=\bar{p}} &= -\frac{M}{\bar{p}^2} + \frac{1}{\bar{p}}\sum_{h=1}^{S} F'_h(1), \qquad (4.21) \\ \frac{\partial H}{\partial p^{h,e}_{t+1}}\bigg|_{p_{t+1}=p_t=\bar{p}} &= -\frac{1}{\bar{p}}F'_h(1), \text{ where} \\ F'_h(1) &= \frac{\sigma_h - 1}{\sigma_h + \varepsilon_h}. \end{aligned}$$

So, $A_h = \frac{\frac{1-\sigma_h}{\sigma_h+\varepsilon_h}}{S+\sum_h\left(\frac{1-\sigma_h}{\sigma_h+\varepsilon_h}\right)}$.[14]

It is possible to show, using the criterion for δ–stability for univariate economies, that this economy is always δ–stable for any $\epsilon_h, \sigma_h > 0$. Consequently, it is E–stable, as well.

Proposition 4.14 *The OLG economy defined above is δ–stable.*

Proof. I have $A_h = \frac{\frac{1-\sigma_h}{\sigma_h+\varepsilon_h}}{S+\sum_h\left(\frac{1-\sigma_h}{\sigma_h+\varepsilon_h}\right)}$ and $\epsilon_h, \sigma_h > 0$. Writing down the second part of the criterion in strict inequalities, I get: $\sum_{h\subseteq(h_1,\ldots,h_p)} A_h < 1 \iff \frac{\sum_{h\in(h_1,\ldots,h_p)} \frac{1-\epsilon_h}{\sigma_h+\varepsilon_h}}{\sum_{h=1}^{S} \frac{\epsilon_h+1}{\sigma_h+\varepsilon_h}} - 1 < 0 \iff$ $\frac{-\sum_{h\notin(h_1,\ldots,h_p)} \frac{1}{\sigma_h+\varepsilon_h} - \sum_{h\in(h_1,\ldots,h_p)} \frac{\epsilon_h}{\sigma_h+\varepsilon_h} - \sum_{h=1}^{S} \frac{1}{\sigma_h+\varepsilon_h}}{\sum_{h=1}^{S} \frac{\epsilon_h+1}{\sigma_h+\varepsilon_h}} < 0$ for any subeconomy $(h_1, \ldots, h_p)$ (including the original economy). For $\epsilon_h, \sigma_h > 0$, the condition is always satisfied. □

The behavior around the steady state equilibrium of the OLG exchange economy considered by Honkapohja and Mitra [43] is presented by the following system of equations:

$$\begin{aligned} \tilde{p}_t &= \sum_{h=1}^{S} A_h \tilde{p}_{t+1}^{h,e}, \text{ where} \qquad (4.22) \\ A_h &= \frac{F_h'(1)}{-\frac{M}{\bar{p}} + \sum_h F_h'(1)}, F_h'(1) = \frac{\omega_{2,h}(2-\rho_h) + \omega_{1,h}\rho_h}{4(\rho_h - 1)}, \frac{M}{\bar{p}} = \frac{1}{2}\sum_h (\omega_{1,h} - \omega_{2,h}), \end{aligned}$$

where $\omega_{1,h}$ and $\omega_{2,h}$ denote the endowment of a single good to the agent of type h for its first and second periods of life, respectively. $\rho_h < 1$ is a parameter of agent of type h born in period t utility function of consumption in the first and the second periods of its life: $U_h(c_{h,t}, c_{h,t+1}) = \left(c_{h,t}^{\rho_h} + c_{h,t+1}^{\rho_h}\right)^{1/\rho_h}$.

Similarly to the OLG economy of the Samuelson type considered by me above, it is possible to show that this economy is always E–stable. Moreover, for the specifications satisfying $\omega_{1,h} > \omega_{2,h}$ for all $h = \overline{1, S}$ (all examples of Honkapohja and Mitra [43] satisfy this specification), the criterion for δ–stability for the univariate economy (Proposition 4.10) allows me to say that the economy is δ–stable.

[14] Notice, that Honkapohja and Mitra [43] do not have minus before the first term in (4.21). My derivations of the reduced form are algebraically analogous to their derivation, and I suspect they lost this minus during derivation. Though their example for values of A_h remains valid for this reduced form, the values of ρ_h's in agents CES utiliy functions in their overlapping generations exchange economy could not be found in a plausible range (that is, $\rho_h < 1$) for their specification of A_h' s.

Proposition 4.15 *The OLG exchange economy (4.22) is E–stable. If $\omega_{1,h} > \omega_{2,h}$ for all $h = \overline{1, S}$, it is δ–stable.*

Proof. To prove the first part of the proposition, I may use only the condition on the parameters of the utility function, $\rho_h < 1$ and the condition that $\omega_{1,h}, \omega_{2,h} \geqslant 0$ From the formula for $F'_h(1)$ I get $\rho_h = \frac{2\omega_{2,h}+4F'_h(1)}{4F'_h(1)+\omega_{2,h}-\omega_{1,h}} < 1$. It leads to inequality $F'_h(1) < \frac{\omega_{1,h}-\omega_{2,h}}{4}$. Using $\frac{M}{\bar{p}} = \frac{1}{2}\sum_{h=1}^{S}(\omega_{1,h} - \omega_{2,h})$, I get $\sum_{h=1}^{S} F'_h(1) < \frac{M}{2\bar{p}}$. Next using the formula for A_h, I get $\sum_{h=1}^{S} F'_h(1) = \frac{\frac{M}{\bar{p}}\sum A_h}{\sum A_h - 1} < \frac{M}{2\bar{p}}$. As $\frac{M}{\bar{p}} > 0$ (from its economic meaning), I arrive at inequality $\frac{\sum A_h + 1}{\sum A_h - 1} < 0$ that leads to $-1 < \sum_{h=1}^{S} A_h < 1$. Condition $\sum_{h=1}^{S} A_h < 1$ is the condition of E–stability. □

To prove the second part of the proposition, I first express A_h via ρ_h,$\omega_{1,h}, \omega_{2,h}$. I have $A_h = \frac{F'_h(1)}{-\frac{M}{P}+\sum_{h=1}^{S} F'_h(1)}$. Substituting for $F'_h(1)$ and $\frac{M}{P}$ I arrive at $A_h = \frac{\frac{\omega_{2,h}+\omega_{1,h}}{1-\rho_h}+\omega_{2,h}-\omega_{1,h}}{\sum_{h=1}^{S}\left(\frac{\omega_{2,h}+\omega_{1,h}}{1-\rho_h}+\omega_{1,h}-\omega_{2,h}\right)}$. Writing down the second part of the criterion in strict inequalities I get:

$$\sum_{h\subseteq(h_1,\ldots,h_p)} A_h < 1 \Longleftrightarrow \frac{-\sum_{h\notin(h_1,\ldots,h_p)}\frac{\omega_{2,h}+\omega_{1,h}}{1-\rho_h} - \sum_{h=1}^{S}\left(\omega_{1,h}-\omega_{2,h}\right) + \sum_{h\in(h_1,\ldots,h_p)}\left(\omega_{2,h}-\omega_{1,h}\right)}{\sum_{h=1}^{S}\left(\frac{\omega_{2,h}+\omega_{1,h}}{1-\rho_h}+\omega_{1,h}-\omega_{2,h}\right)} <$$

0 for any subeconomy $(h_1, ..., h_p)$, including the original economy. Since we have $\rho_h < 1, \omega_{1,h} > \omega_{2,h} \geq 0$, the condition is always satisfied. *Q.E.D.*

In addition, I will show how my criterion works for the specification of the reduced form used by Honkapohja and Mitra [43]. Let us say that these values of A_h are possible for some other model. Honkapohja and Mitra [43] consider the following specifications: $S = 3$, $A_1 = 0.1$, $A_2 = -0.2$ and $A_3 = -0.5$; $S = 3$, $A_1 = -15$, $A_2 = 0.5$ and $A_3 = 0.6$; $S = 3$, $A_1 = -15$, $A_2 = 1.1$ and $A_3 = 0.6$.

Since I have been able to derive a criterion for δ–stability in the univariate case, I can say, looking only at the structure of the model, whether it is stable under all types of heterogeneous learning, or not, without looking for examples with various degrees of inertia of agents that violate convergence.

For the first specification, applying the criterion for δ–stability in the univariate case, and finding that the same sign condition is violated, one is left to check the condition for stability of subeconomies: since the setup here is non–stochastic, one is left to check that all average economies corresponding to subeconomies are not unstable and at least one of them is stable, and it can be easily checked by considering stability of the corresponding average economies. $A_1 = 0.1 < 1, A_2 = -0.2, A_3 = -0.5 < 1, A_1 + A_2 = -0.1 < 1, A_1 + A_3 = -0.4 < 1, A_2 + A_3 = -0.7 < 1, A_1 + A_2 + A_3 = -0.6 < 1$. So, all average economies corresponding to all subeconomies (including the original economy) are E–stable. This means that economy is δ–stable.

For the second specification, using the criterion above, it is clear that the economy will not be δ–stable, as none of the conditions of the criterion is satisfied: the same sign condition is violated, and there exists an average economy corresponding to subeconomy $(2,3)$ for which $A_{(2,3)} = A_2 + A_3 = 1.1 > 1$, that is, this average economy is E–unstable. Alternatively, one can easily check that eigenvalues of $\Omega_{(2,3)} = \begin{pmatrix} A_2 - 1 & A_3 \\ A_2 & A_3 - 1 \end{pmatrix} = \begin{pmatrix} -0.5 & 0.6 \\ 0.5 & -0.4 \end{pmatrix}$ are -1 and 0.1, which violates the stability conditions.

A similar situation is for the third specification. The economy is not δ–stable since the conditions of the criterion are not satisfied: the same sign condition is violated, as A_h's have different signs, and there exist average economies corresponding to subeconomies (2) and $(2,3)$ for which $A_{(2)} = A_2 = 1.1 > 1$ and $A_{(2,3)} = A_2 + A_3 = 1.7 > 1$, that is, these average economies are E–unstable. Alternatively, one can check that eigenvalues of $\Omega_{(2,3)} = \begin{pmatrix} A_2 - 1 & A_3 \\ A_2 & A_3 - 1 \end{pmatrix} = \begin{pmatrix} 0.1 & 0.6 \\ 1.1 & -0.4 \end{pmatrix}$ and of $\Omega_{(2)} = A_2 - 1 = 0.1$ are -1 and 0.7, and 0.1, respectively, which violates the stability conditions.

In order to further demonstrate the power of the derived criterion for δ–stability, I will consider the case of more than 3 agents in the economy. Let us consider $S = 6$, $A_1 = -0.1,\ A_2 = -0.2,\ A_3 = -0.5, A_4 = -15,\ A_5 = 0.5,\ A_6 = 0.5$.

The economy under this specification is δ–stable, notwithstanding that the same sign condition is violated and there is an average economy corresponding to subeconomy $(5,6)$ for which $A_{(5,6)} = A_5 + A_6 = 1$. The condition of the criterion is satisfied. Indeed, all $A_{(h_1,\ldots,h_p)} = \sum\limits_{(h_1,\ldots,h_p)} A_h$ are less or equal than 1, and for each size p there exists an average economy for which this coefficient is strictly less than one: for $p = 1, A_{(1)} = -0.1 < 1$, for $p = 2, A_{(1,2)} = -0.3 < 1$, for $p = 3, A_{(1,2,3)} = -0.8 < 1$, for $p = 4, A_{(1,2,3,4)} = -15.8 < 1$, for $p = 5, A_{(1,2,3,4,5)} = -15.3 < 1$, for $p = 6, A_{(1,2,3,4,5,6)} = -14.8 < 1$. So, even if the economy contains a subeconomy which is E–unstable (lies on the boundary of the stability/instability) and the same sign condition is violated, the whole economy can be

δ–stable.

4.5.2 Multivariate Case

I demonstrate the aggregate economy sufficient conditions on a model of simultaneous markets with structural heterogeneity[15]. The idea is to add more economic interpretation to these conditions on an example of a particular multivariate model.[16]

The economic environment is given by the following equation:

$$p_t = l + vd_t + \varepsilon_t,$$

which is the demand function in matrix form for different goods $j = \overline{1, J}$.

p_t is a $J \times 1$ vector of prices, which are endogenous variables in this model, l is a vector of intercepts, v is a $J \times J$ matrix which corresponds to the inverse of the matrix of price effects. $d(t)$ is a vector of quantities of the J goods, $\varepsilon_{j,t} = f_j \varepsilon_{j,t-1} + v_{j,t}$, $\varepsilon_{j,t}$ are demand shocks, $|f_j| < 1$, and $v_{j,t}$ are independent white noise processes.

There are S types of suppliers with supply functions:

$$s_t^h = g^h + n^h \hat{E}_{t-1}^h p_t, h = \overline{1, S},$$

which depend on the expected price due to a production lag. Each supplier produces all J goods. $s(h, t)$ is a $J \times 1$ vector of goods supplied by type h supplier.

It is further assumed that different outputs are produced in independent processes by each producer h, so n^h is a positive diagonal matrix. Expectations (non-rational, in general) of prices are formed by each supplier at the end of period $t-1$ before the realization of the demand shock $\varepsilon_{t.}$

The market clearing condition, $d_t = \sum\limits_{h=1}^{S} s_t^h$, leads to the following reduced form:

$$p_t = l + v \left(\sum_{i=1}^{S} g^i \right) + \sum_{h=1}^{S} v n^h \hat{E}_{t-1}^h p_t + \varepsilon_t.$$

For the case with equal weights of single agent types used in calculating aggregate expectations, the aggregate stability sufficient condition for this model has the form

$$\sum_i \psi_i |v_{ij}| < \frac{\psi_j}{S n_{jj}^h}, \forall j, h.$$

[15] The author expresses sincere thanks to Seppo Honkapohja who suggested to use this example.

[16] δ–stability of a bivariate (New Keynesian) model under two types of optimal monetary policy rules of a policy maker is considered in a companion paper (Bogomolova and Kolyuzhnov [6]).

This condition can be derived by the direct application of Proposition 4.3 to the given model.

I am going to show now that this sufficient condition for δ–stability at the same time is a sufficient condition for E–stability of the aggregate (univariate) cobweb model. In order to show this, I have to derive, first, the aggregate supply and demand curves using weights of aggregation across agents and expectations I used to derive the sufficient conditions for δ–stability above.

So, the aggregate demand curve for the price index[17] can be derived as follows:

$$P_t = \sum_i \psi_i p_{it} = \left(\sum_i \psi_i v_{i1}\right) d_{1t} + ... + \left(\sum_i \psi_i v_{iJ}\right) d_{Jt} + \sum_i \psi_i l_i + \sum_i \psi_i \varepsilon_{it} <$$

$$< \left(\sum_i \psi_i \left|v_{i1}\right|\right) d_{1t} + ... + \left(\sum_i \psi_i \left|v_{iJ}\right|\right) d_{Jt} + \sum_i \psi_i l_i + \sum_i \psi_i \varepsilon_{it} =$$

$$= \underbrace{\left(\sum_i \psi_i \left|v_{i1}\right| + ... + \sum_i \psi_i \left|v_{iJ}\right|\right)}_{r_p} \underbrace{\left\{\frac{\left(\sum_i \psi_i |v_{i1}|\right) d_{1t} + ... + \left(\sum_i \psi_i |v_{iJ}|\right) d_{Jt}}{\sum_i \psi_i |v_{i1}| + ... + \sum_i \psi_i |v_{iJ}|}\right\}}_{D^{AG}} + \sum_i \psi_i l_i + \sum_i \psi_i \varepsilon_{it} =$$

$$= r_p D^{AG} + \sum_i \psi_i l_i + \sum_i \psi_i \varepsilon_{it}.$$

Note that here, aggregating over the elements of the price vector, I obtain the demand function in terms of the price index. This is an example of economic interpretation of the aggregation procedure that I propose in my paper, in particular, of assigning weights to the endogenous variables.

To derive the aggregate supply curve for the price index I, first, write the aggregate (over all supplier types) supply equation:

$$\sum_h s_t^h = \sum_h g^h + \sum_h n^h \hat{E}_{t-1}^h p_t = \sum_h g^h + \left(\sum_h n^h\right) \hat{E}_{t-1}^{AG} p_t.$$

Then I write equations for each component of the supply vector: the aggregate supply of each product equations. So, for each product j,

$$\sum_h s_{jt}^h = \sum_h g_j^h + \left(\sum_h n^h \hat{E}_{t-1}^h p_t\right)_j = \sum_h g_j^h + \left(n_{11}^1 + ... + n_{JJ}^S\right) \hat{E}_{t-1}^{aggreg} p_t^j.$$

Next, I aggregate over all supply equations using weights ψ_j. Aggregating across endogenous variables (prices) to get the price index, I finally get the aggregate supply curve for

[17] To get this function, I aggregate the individual demand functions, not the reduced form equations (in which case I would obtain an equation for the intertemporal equilibrium price index).

the aggregate model

$$\hat{E}_{t-1}^{AG}P_{t}. = \left(\frac{\psi_1}{n_{11}^1+...n_{11}^S}\right)\sum_h s_{1t}^h + ... + \left(\frac{\psi_J}{n_{JJ}^1+...n_{JJ}^S}\right)\sum_h s_{Jt}^h - \sum_j \psi_j \left(\sum_h g_j^h / \left(n_{11}^1 + ... + n_{JJ}^S\right)\right) =$$

$$= \underbrace{\left(\frac{\psi_1}{n_{11}^1+...n_{11}^S} + ... + \frac{\psi_J}{n_{JJ}^1+...n_{JJ}^S}\right)}_{r_m} \underbrace{\left\{\frac{\left(\frac{\psi_1}{n_{11}^1+...n_{11}^S}\right)\sum_h s_{1t}^h + ... + \left(\frac{\psi_J}{n_{JJ}^1+...n_{JJ}^S}\right)\sum_h s_{Jt}^h}{\left(\frac{\psi_1}{n_{11}^1+...n_{11}^S}\right) + ... + \left(\frac{\psi_J}{n_{JJ}^1+...n_{JJ}^S}\right)}\right\}}_{S^{AG}} -$$

$$- \sum_j \psi_j \left(\sum_h g_j^h / \left(n_{11}^1 + ... + n_{JJ}^S\right)\right) =$$

$$= r_m S^{AG} - \sum_j \psi_j \left(\sum_h g_j^h / \left(n_{11}^1 + ... + n_{JJ}^S\right)\right).$$

Thus, we have the following aggregate cobweb model in structural form:

$$P_t = r_p D^{AG} + \sum_i \psi_i l_i + \sum_i \psi_i \varepsilon_{it} \text{ is the aggregate demand curve}$$

$$\hat{E}_{t-1}^{AG}P_t = r_m S^{AG} - \sum_j \psi_j \left(\sum_h g_j^h / \left(n_{11}^1 + ... + n_{JJ}^S\right)\right) \text{ is the aggregate supply curve,}$$

where

$$\begin{aligned} r_p &= \sum_i \psi_i |v_{i1}| + ... + \sum_i \psi_i |v_{iJ}| \\ r_m &= \left(\frac{\psi_1}{n_{11}^1+...n_{11}^S} + ... + \frac{\psi_J}{n_{JJ}^1+...n_{JJ}^S}\right). \end{aligned}$$

It is clear that from the sufficient condition for δ–stability $\sum_i \psi_i |v_{ij}| < \frac{\psi_j}{Sn_{jj}^h}, \forall j, h$, follows $\sum_i \psi_i |v_{ij}| < \frac{\psi_j}{S\max_h\{n_{jj}^h\}}, \forall j$ and, in turn, $\frac{\psi_j}{S\max_h\{n_{jj}^h\}} < \frac{\psi_j}{n_{jj}^1+...n_{jj}^S}, \forall j$. Thus, the sufficient condition for δ–stability in this class of models, $\sum_i \psi_i |v_{ij}| < \frac{\psi_j}{Sn_{jj}^h}, \forall j, h$, is the condition for E–stability of the aggregate cobweb model ($r_m > r_p$).

4.6 Conclusion

My paper to some extent resolves the open question posed by Honkapohja and Mitra [43]. As has been mentioned, Honkapohja and Mitra [43] provide a general stability condition (criterion) for the case of persistently heterogeneous learning — a joint restriction on matrices of structural parameters and degrees of inertia, which implies that stability in such an economy is determined by the interaction of structural heterogeneity and learning heterogeneity. For the general (multivariate) case, however, it was not possible to derive easily interpretable stability conditions expressed in terms of an economy aggregated only across agent types. Honkapohja and Mitra [43] have derived sufficient conditions in terms of the structure of the economy, but this condition is very general: it requires D–stability and H–stability of the structural matrices.

In this paper, I attempt to fill this gap and provide easily interpretable sufficient and necessary conditions for such a stability. Based on the analysis of the negative diagonal dominance, the alternative definition of D–stability, and the characteristic equation analysis, I have been able to derive two groups of sufficient conditions and one group of necessary conditions for δ–stability, that is, stability under heterogeneous learning, independent of heterogeneity in parameters of learning algorithms. I have found an easily interpretable unifying condition which is sufficient for convergence of an economy under mixed RLS/SG learning with different degrees of inertia towards a rational expectations equilibrium for a broad class of economic models and a criterion for such a convergence in the univariate case. The conditions are formulated using the concept of a subeconomy and a suitably defined aggregate economy.

In particular, using the negative diagonal dominance (sufficient for D–stability) and my concept of aggregating an economy (both across agent types and endogenous variables), I have obtained sufficient conditions for δ–stability expressed in terms of E–stability of the aggregate economy and its structure. These were summarized as the aggregate economy sufficient conditions. One of them can serve as a rule of thumb for checking a model for δ–stability.

I have found a unifying condition for the most general case of heterogeneous learning in linear forward–looking models. Though it is quite restrictive, my main achievement was to show that such a simple condition with the E–stability meaning of some aggregate economy (a notion that has already proved useful as a condition for stability under heterogeneous learning in previous learning literature) does exist for a large class of models. The economic example provided in the end of the paper demonstrates the application of the aggregate economy conditions.

Next, based on the analysis of the alternative definition of D–stability, I have obtained sufficient conditions on the structure of the economy summarized as the "same sign" conditions. Further, based on the analysis of the characteristic equation and the requirement for negativity of all eigenvalues (necessary and sufficient for stability), I have derived a group of necessary conditions. Their failure can be used as an indicator of non–δ–stability.

Moreover, using the alternative definition of D–stability and the characteristic equation approaches, I obtain the criterion for δ–stability in the univariate case. On the example of two types of OLG models I show that this criterion can be easily used to test an economy for δ–stability.

Chapter 5

Optimal Monetary Policy Rules: The Problem of Stability Under Heterogeneous Learning

Optimal Monetary Policy Rules: The Problem of Stability Under Heterogeneous Learning*

Anna Bogomolova and Dmitri Kolyuzhnov†

CERGE–EI‡
Politických vězňů 7, 111 21 Praha 1,
Czech Republic

Abstract

In this paper we extend the analysis of optimal monetary policy rules in terms of stability of an economy, started by Evans and Honkapohja [31], to the case of heterogeneous private agents learning. Following Giannitsarou [37], we pose the question about the applicability of the representative agent hypothesis to learning. This hypothesis was widely used in learning literature at early stages to demonstrate convergence of an economic system under adaptive learning of agents to one of the rational expectations equilibria in the economy. We test these monetary policy rules in the general setup of the New Keynesian model that is a work horse of monetary policy models today. It is of interest to see that the results obtained by Evans and Honkapohja [31] for the homogeneous learning case are replicated for the case when the representative agent hypothesis is lifted.

JEL Classification: C62, D83, E31, E52
Keywords: monetary policy rules, New Keynesian model, adaptive learning, stability of equilibrium, heterogeneous agents

---

*The major part of this paper was written while both authors stayed as visiting Ph.D. students at the University of Cambridge, UK. The authors express special thanks to their supervisor while at Cambridge, Seppo Honkapohja, for continuous support of this research and immeasurable help and advice. All errors are the authors' responsibility.

The authors express thanks to the participants of the European Economics and Finance Society (EEFS 2006) 5th Annual Conference "European Labour Markets, Trade and Financial Flows and Economic Growth" in Heraklion, Crete, Greece May, 18-21, 2006 and the conference of the Society for Economic Dynamics (SED 2007), Prague, Czech Republic, June, 2007, for valuable discussion.

The paper won the first place in the Czech Econometric Society competition for the best Ph.D. student paper.

† *{Anna.Bogomolova, Dmitri.Kolyuzhnov }@cerge-ei.cz.*

‡CERGE–EI is a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic.

5.1 Introduction

The stabilization monetary policy design problem is very often studied in the New Keynesian model. Using the environment of this model, we may study different monetary policy rules to find out which is more efficient in smoothing business cycle fluctuations and also which monetary policy rule would not lead to indeterminacy of equilibria in our model. For a comprehensive overview of various interest rate rules in the New Keynesian model, one can address Woodford [69]. Also, very often cited works on monetary policy design are Clarida, Gali, and Gertler [19, 20]. Svensson [62] gives a clear distinction between instrument and target rules and implications of their use.

A number of recent studies also consider the New Keynesian model environment with adaptive learning of agents. Examples are works of Evans and Honkapohja [30, 31], Bullard and Mitra [11] and Honkapohja and Mitra [42] on stability of an economy under various policy rules. Evans and Honkapohja [30, 31] take up the issue of stability under learning for optimal monetary policies in economies with adaptive learning.

The concept of adaptive learning of agents in economic models is introduced as a specific form of bounded rationality advocated by Sargent [60]. According to the argument of Sargent [60], it is more natural to assume that agents face the same limitations economists face (in a sense that economists have to learn the model structure and its parameter values themselves) and view agents as econometricians when forecasting the future state of the economy.

Using adaptive learning in an economy makes it possible to test the validity of the rational expectations hypothesis by checking if a given dynamic model converges over time to the rational expectations equilibrium (REE) implied by the model. It can also be used as a selection device in models with multiple equilibria. Even if the model has a unique REE, it is still of interest to see if the rational expectations (RE) hypothesis holds under learning, which is done by checking if our model under learning converges to a given REE. In both cases (multiple or unique REE), one has to check certain stability conditions. After this analysis of stability conditions, the next step could be studying policy rules for effectiveness and indeterminacy, assuming or making sure that the stability conditions on the model structure are satisfied.

That is why, before we start analyzing particular monetary policies for efficiency (evaluating a particular type of policy: Taylor rule, optimization–based rule with or without commitment), we should take a general type of a linear policy feedback rule, plug it into our structural form of the New Keynesian model and obtain some general linear reduced form (RF) of this model. All things being equal (the same structural equations: Phillips and IS curves), we can obtain different RFs depending on the policy rule used by the policy maker. Hence, we obtain different REEs and different stability results. Then we should study a given reduced form for stability in order to see if a given REE is chosen. In this paper, we study the stability of a New Keynesian model under the following classification of policy rules introduced by Evans and Honkapohja [31].

Depending on the assumptions of the central bank about the expectations of the private agents (firms, households), Evans and Honkapohja [31] divide all policy rules into fundamentals–based rules and expectations–based rules. The fundamentals–based rule is obtained if the policy maker assumes RE of private agents, while the expectations–based rule takes into account possibly non–rational expectations of agents (assuming that these expectations are observable to the central bank).[1]

We consider the stability question under the assumption of heterogeneous learning of agents. As has been shown in Giannitsarou [37] and Honkapohja and Mitra [43], stability results may be different under homogeneous and heterogeneous learning. Honkapohja and Mitra [43] also demonstrate that stability may depend on the interaction of structural heterogeneity and learning heterogeneity, and Honkapohja and Mitra [42] examine how structural heterogeneity in the New Keynesian model may affect stability results under various types of policy rules.

Note that though Honkapohja and Mitra [42] consider heterogeneity in learning in the New Keynesian model, their definition of heterogeneity implies a situation when the central bank and private agents have (possibly) different learning algorithms with (possibly) different parameters of these algorithms. They essentially consider the situation when all private agents could be considered as one representative agent, and in this sense learning of private agents considered by Honkapohja and Mitra [42] is homogeneous. In some sense, the situation considered by Honkapohja and Mitra [42] could be called two-sided learning in a structurally heterogeneous bivariate economy.

[1] We should note here that in Taylor–type rules the current value of interest rate depends on the current values of inflation and output gap. In this paper we study stability under feedback rules that are derived from the policy maker minimization problem, in particular, study their two categories, according to Evans and Honkapohja [31]: fundamentals–based and expectations–based. Stability under Taylor–type rules, which do not fall under this classification, will be studied later in a separate work.

In this paper we do not consider learning of the central bank and assume, following Evans and Honkapohja [31], that the policy maker takes expectations of private agents as given or assumes and knows the exact structure of their rational expectations; at the same time we fully exploit the case when private agents have heterogeneous learning. The case of the internal central bank forecasting (that includes Taylor rules) in a situation of heterogeneous learning of private agents, which develops the model of Honkapohja and Mitra [42] since Honkapohja and Mitra [42] consider only the situation of a representative private agent, is the topic of our further research.

It turns out that under the fundamentals–based linear feedback policy rule (optimization–based), learning in our model never converges to the REE of the model. Evans and Honkapohja [31] demonstrate this instability result for the homogeneous recursive least squares (RLS) and for the stochastic gradient (SG) learning,[2] while we obtain a similar instability result for the three types of heterogeneous learning considered by Giannitsarou [37].

The other category of policy rules — expectations–based rules — is supposed to react to agents' expectations. Under certain conditions, we can have stability under such rules. Evans and Honkapohja [31] obtain a stability result for homogeneous RLS or for SG learning. We obtain a stability result (with conditions on the model structure) for the case of the three types of heterogeneous learning considered by Giannitsarou [37].

Originally, when heterogeneous learning in a general setup of self-referential linear stochastic models was studied by Giannitsarou [37], the purpose of introducing heterogeneous learning of agents was to see if the representative agents hypothesis influences stability results, i.e., if one may always apply this hypothesis. For some cases, it is demonstrated that it does make sense to consider the heterogeneous setup. Our paper is about stability under monetary policy rules, so, though we, in fact, prove that the representative agent hypothesis holds true for the New Keynesian model, the accent of our paper is shifted away from testing the importance (influence) of the representative agent hypothesis.

We, essentially, apply the stability analysis of the model under heterogeneous

[2]Honkapohja and Mitra [43] and we in this paper consider two possible algorithms used to reflect bounded rationality of agents: RLS and SG learning algorithms (which are examples of econometric learning). Their description can be found, e.g., in Evans and Honkapohja [29], Honkapohja and Mitra [43], Giannitsarou [37], and Evans, Honkapohja and Williams [32]. Both are used by agents to update the estimates of the model parameters. Essentially, the difference is as follows. The RLS algorithm has two updating equations: one—for updating parameters entering the forecast functions, the other—for updating the second moments matrix (of the model state variables). The SG algorithm assumes this matrix fixed.

learning in the same manner the stability analysis of the model under homogeneous (when all agents can be substituted with a representative agent) learning is applied in Evans and Honkapohja [31].[3] In our paper, we link the study of stability conditions under a certain category of linear monetary policy rules of [31] with the study of stability under heterogeneous learning of Giannitsarou [37].

We first show that in the New Keynesian–type of models, stability can be analyzed using the structural parameters, whatever the type of heterogeneous learning, using the general criterion of Honkapohja and Mitra [43]. These results are the structural matrix eigenvalues sufficient and necessary conditions for stability of a structurally homogeneous model derived in this paper and the aggregate economy sufficient conditions derived in Kolyuzhnov [51], where the concept of stability under heterogeneous learning, termed as δ–stability, is introduced. Then we apply these results to derive stability and instability results under heterogeneous learning for the two categories of feedback rules: fundamentals–based and expectations–based, in the model with an arbitrary number of agent types.

Summarizing all the above, our work now looks, on the one hand, like a link between the study of stability under monetary policy rules for homogeneous learning of Evans and Honkapohja [31] and the study of stability conditions under heterogeneous learning of Giannitsarou [37], — the link through the δ–stability conditions derived by us for the general setup of Honkapohja and Mitra [43] and through the general stability criterion of Honkapohja and Mitra [43]. On the other hand, this study can serve as one more economic example demonstrating the application of δ–stability sufficient and necessary conditions.

The structure of the paper is as follows. In the next section we present the basic New Keynesian model. In Section 3 we discuss the general stability results under heterogeneous learning and the concept of δ–stability introduced in Kolyuzhnov [51]. In Section 4, we give necessary and sufficient conditions for δ–stability for structurally homogeneous

[3] Evans and Honkapohja [31] study stability conditions under monetary policy rules for the case of homogeneous learning. Their major input is (both for the one–sided learning and the two–sided learning) to have shown that under fundamentals based rules the REE of the model is always unstable, while under the expectations based rule there is always stability. In the two cases the reduced form of the model is different, which has, as a consequence, the difference in the stability results. So, the policy implication of such a stability analysis is that, given the structure of the model (the two structural New Keynesian equations), the central bank can influence (determine) the outcome of its policy by selecting the appropriate optimal monetary policy: the one that guarantees convergence to a particular REE.

models. Section 5 describes the two types of optimal policy rules and the structure of the reduced forms under each type. In Section 6 we provide stability and instability results for the types of optimal monetary policies considered in application to the New Keynesian model. Section 7 concludes.

5.2 Model

The model that we consider is a general New Keynesian model with observable stationary AR(1) shocks. The structural form of the model looks as follows

$$x_t = c_1 - \phi\left(i_t - \widehat{E}_t\pi_{t+1}\right) + \widehat{E}_t x_{t+1} + \chi_1' w_t \tag{5.1}$$

$$\pi_t = c_2 + \lambda x_t + \beta\widehat{E}_t\pi_{t+1} + \chi_2' w_t, \tag{5.2}$$

where the first equation is for the IS curve and the second equation is for the Phillips curve. $w_t = \left[\begin{array}{ccc} w_{1t} & ... & w_{kt} \end{array}\right]'$ is a vector of observable AR(1) shocks[4],

$$w_{it} = \rho_i w_{it-1} + \nu_{it}, |\rho_i| < 1, \nu_{it} \sim iid\left(0, \sigma_i^2\right), i = \overline{1,k} \tag{5.3}$$

To introduce heterogeneity in the model, we assume that we have S types of private agents characterized by their share $\zeta_h > 0$ in the economy, $\sum_{h=1}^{S}\zeta_h = 1$. So, $\widehat{E}_t x_{t+1} = \sum_{h=1}^{S}\zeta_h\widehat{E}_t^h x_{t+1}$, $\widehat{E}_t\pi_{t+1} = \sum_{h=1}^{S}\zeta_h\widehat{E}_t^h\pi_{t+1}$, where $\widehat{E}_t^h x_{t+1}$ and $\widehat{E}_t^h\pi_{t+1}$ are expectations (in

[4]Typically, New Keynesian models include only an observable component, which is assumed to follow an AR(1) process. However, there are specifications including both observable and unobservable shocks. For example, Evans and Honkapohja [32], who study stability rules under recursive least squares learning, include unobservable shocks to the New Keynesian model equations. In our case a more general specification with unobservable shocks would contain additional term $\Omega_1\epsilon_t$ in the IS curve and $\Omega_2\epsilon_t$ in the Phillips curve, where $\epsilon_t = \left[\begin{array}{ccc} \epsilon_{1_t} & ... & \epsilon_{m_t} \end{array}\right]'$ are unobservable shocks, $\epsilon_{i_t} \sim iid\left(0, \gamma_i^2\right)$, $i = 1, ..., m$, not correlated with observable shocks g_t.

Of course, these unobservables do not bring a difference into the stability results (that is why we omit them in the model analyzed), but introducing them into the setup has its own reasoning. For example, it makes sense to introduce unobservable shocks into structural equations when we consider central bank learning structural coefficients of the model. If we have only observable shocks (which play a role of just another regressor – some exogenous variable) as well as other observable regressors, we will evaluate the equations' coefficients exactly if we have a sufficient number of observations. In this case learning is trivial: the convergence will be very quick if initially we did not have enough observations, but gained them over a short period of time.

If we think of how these unobservable shocks can emerge at the micro foundations level, we may think of the following economic interpretation. For example, let us assume that preference and technology shocks consist of observable and unobservable components. As for preference shocks, we can imagine a qualitative change in our preferences, such that we know how the shock has changed our preferences qualitatively, but we cannot precisely measure this change quantitatively. A similar interpretation can be given to the technological shock. What we have measured enters as an observable component, while the measurement error (which always exists since we assume that our quantitative measurement of any change is imprecise) is treated as an unobservable component.

general, non–rational) of private agent of type h made at time t about the next period output gap and inflation, respectively.

The model (5.1), (5.2) and (5.3) is a general formulation of models derived from microfoundations that are considered in macroeconomics and monetary economics literature. The two basic equations of the New Keynesian model, which are the Phillips curve and the IS curve are derived from the optimal problems of the representative household and the representative monopolistically competitive firm, with the assumption of Calvo [13] pricing mechanism in the firms' price–setting decision. So the two New Keynesian curves are derived using the optimality conditions of the private agents (households and firms). The derivation of these two curves for the standard New Keynesian model setup can be found, e.g., in Walsh [63]. The description of the New Keynesian model can also be found in Woodford [68, 69] and in Christiano, Eichenbaum, and Evans [18].

In solving their optimization problems, private agents are assumed to take the interest rate (entering the IS curve equation) as given. The interest rate, in turn, is set by the policy maker — the central bank. In various studies of monetary policy issues (in the New Keynesian framework), it is normally assumed that the policy maker uses some linear feedback rule to set the interest rate. In general, a feedback rule that is derived from the loss function minimization problem determines how the interest rate reacts to the expected values of the model's endogenous variables (inflation and output gap in the New Keynesian model) and the model's exogenous variables (various shocks, e.g., technology shock, preference shock, cost-push shock). Instrument rules, like Taylor–type rules, are designed to respond to the target variables (e.g., inflation and output gap). As is noted in the introduction, Taylor–type rules will be considered in a separate study.

Plugging the feedback rule into the IS curve equation, we obtain the model reduced form. Using the same New Keynesian equations (IS and Phillips curves), we can obtain different reduced forms for different policy rules, i.e. other things being equal, the reduced form structure depends on the policy rule. It depends not only on the type of it (Taylor or optimization–based), but, as is demonstrated by Evans and Honkapohja [31], on the assumption of the central bank about private agents expectations, resulting either in the fundamentals–based or in the expectations–based category of feedback rules.

After plugging some monetary policy rule of the central bank i_t, assuming that the central bank knows expectations of private agents or assumes and knows the form of rational expectations of agents (we will talk about the types of optimal monetary policy rules later), the model can be written in the reduced form that has a general representation of a bivariate system with a stationary AR(1) observable shocks process

$$y_t = \alpha + A\hat{E}_t y_{t+1} + Bw_t, \quad (5.4)$$

$$y_t = \begin{bmatrix} \pi_t & x_t \end{bmatrix}' \quad (5.5)$$

and (5.3).

In what follows, for the derivation of our stability results we may allow for some generalization (as it is just a matter of notation compared to the bivariate model) and consider a multivariate (not just bivariate) system (5.4) with a stationary AR(1) observable shocks process (5.3).

In our notation, the reduced form is written in such a way that it includes all factors that appear in the structural form. This means that the absence of some factor in the reduced form in our notation is expressed by the corresponding zero column of matrix B. Note that here we adopt such a notation in order to be able later to consider different specifications of learning algorithms that include factors from different sets.[5] So our notation is the most general that can be.

In adaptive learning models of bounded rationality it is assumed that agents do not know the rational expectations equilibrium and instead have their own understanding of the relation between variables in the model. The coefficients in this relation (that are called beliefs) are updated each period as new information on observed variables arrives (in this respect agents are modeled as if they were statisticians, or econometricians) For the beginning, we assume that agents have the following perceived relation among the variables in the economy, which is called the perceived law of motion (PLM)

$$y_t = a^h + \Gamma^h w_t,$$

with $a^h = \begin{bmatrix} a_1^h & a_2^h \end{bmatrix}', \Gamma^h = \begin{bmatrix} \gamma_{11}^h & \gamma_{12}^h & \cdots & \gamma_{1k}^h \\ \gamma_{21}^h & \gamma_{22}^h & \cdots & \gamma_{2k}^h \end{bmatrix}$ in the bivariate case,

that includes all components of w_t. A bit later we weaken this assumption. Though we assume that the parameters of the PLM may differ across agents, we assume that the

[5] An example when a model reduced form may not include all shocks that are present as factors in the model structural form can be found in Evans and Honkapohja [31], who used the New Keynesian model setup of Clarida, Gali, Gertler [19].

structure of the PLMs is the same for all agents. We may also write the average (or aggregate) PLM using the weights of agents.

$$y_t = a + \Gamma w_t, \text{ where } a = \sum_{h=1}^{S} \zeta_h a^h, \Gamma = \sum_{h=1}^{S} \zeta_h \Gamma^h. \tag{5.6}$$

Thus agents have the following forecast functions based on their PLMs

$$\widehat{E}_t^h y_{t+1} = a^h + \Gamma^h diag(\rho_1, ..., \rho_k) w_t$$

and consequently the average forecast function is given by

$$\widehat{E}_t y_{t+1} = \sum_{h=1}^{S} \zeta_h \left(a^h + \Gamma^h diag(\rho_1, ..., \rho_k) w_t \right) = a + \Gamma diag(\rho_1, ..., \rho_k) w_t. \tag{5.7}$$

After plugging the average forecast function (5.7) corresponding to the average PLM (5.6) into the reduced form (5.4), we derive the actual law of motion (ALM)

$$y_t = Aa + \alpha + \left(A\Gamma diag(\rho_1, ..., \rho_k) + B w_t \right). \tag{5.8}$$

The rational expectations equilibrium (REE) defined as $E_t y_{t+1} = \widehat{E}_t y_{t+1} = \widehat{E}_t^i y_{t+1}$ (see, e.g., Sargent [60] or Evans and Honkapohja [29] for the meaning of the RE concept) can be calculated by equating the parameters of the average PLM (5.6) with the corresponding parameters of the ALM (5.8). If we define the T–map as a mapping of beliefs from the average PLM (5.6) to the ALM (5.8),

$$T(a, \Gamma) = \left(Aa + \alpha, A\Gamma diag(\rho_1, ..., \rho_k) + B \right), \tag{5.9}$$

we will be able to write the REE condition as $T(a, \Gamma) = (a, \Gamma)$.

Now we will widen the set of PLMs considered. Let us start with the following definition.

Definition 5.1 *The active factors set is a subset of a set of histories of* w_{i_t} *up to time* t *and a constant used by agents in their PLMs.*[6]

Following the definition, we renumber the subscripts corresponding to regressors that are included into agents' active factors set from 1 to k', and denote the set of subscripts taken from $\{1, ..., k\}$ corresponding to the active factors set as $\widetilde{I}$. Assuming, as before,

[6]Note that by the active factors set we mean not the variables that agents are actually aware of at time t, but essentially those that are used by agents in their PLMs (a subset that may be smaller than the subset of actually available variables).

that all agents have the same structure of their individual PLMs, agents now are assumed to have the following average perceived law of motion (PLM)

$$y_t = a + \widetilde{\Gamma}\widetilde{w}_t$$

$$\text{with } a = \begin{bmatrix} a_1 & a_2 \end{bmatrix}', \widetilde{\Gamma} = \begin{bmatrix} \widetilde{\gamma}_{11} & \widetilde{\gamma}_{12} & \cdots & \widetilde{\gamma}_{1k'} \\ \widetilde{\gamma}_{21} & \widetilde{\gamma}_{22} & \cdots & \widetilde{\gamma}_{2k'} \end{bmatrix} \text{ in the bivariate case,}$$

where $\widetilde{w}_t$ consists of the components of w_t included in agents' active factors set. Consequently, T–map (5.9) can be rewritten as

$$\widetilde{T}(a, \widetilde{\Gamma}) = \left(Aa + \alpha, A\widetilde{\Gamma} diag(\rho_1, ..., \rho_k) + \widetilde{B}\right).$$

where $\widetilde{B}$ consists of columns of matrix B that correspond to the active factors set.

Similarly, one may try to write the REE condition as $\widetilde{T}(a, \widetilde{\Gamma}) = (a, \widetilde{\Gamma})$. However, in this case, it is clear that for the existence of a REE, agents have to include into their active factors set those factors w_{i_t} that correspond to non-zero columns of matrix B in the reduced form. A PLM which consists only of such factors is a PLM that corresponds to the so–called minimal state variable (MSV) solution. Also, in the above PLMs we have used the following assumption.

Assumption 5.1 *Agents include in their PLM of each endogenous variable all factors from their active factors set.*[7]

Essentially, Assumption 5.1 postulates that we may write each agent's PLM equations in matrix form, without a priori setting coefficients at some factors to zero. In addition, we assume that all agents use the same set of factors (which in matrix form means that they use the same vector). We also note here that a similar assumption on the matrix formulation of PLMs has been made by Giannitsarou [37] and Honkapohja and Mitra [43].[8]

[7] So we exclude situations when agents do not include into the PLM equation of one endogenous variable some factor having a zero coefficient in matrix B of the reduced form, while including the same factor in the PLM equation of the other endogenous variable, with this factor having a non–zero coefficient in matrix B of the reduced form. We assume that agents do not know the true structure of the reduced form and use all the available information to form their expectations. So, if one factor is present in one PLM equation, it is present in another PLM equation.

[8] Notice that here we also do not consider situations of the restricted perceptions equilibrium (RPE),

The Propositions below state the necessary and sufficient conditions for the existence and for the uniqueness of a REE in a general multivariate model with stationary AR(1) observable shocks. These conditions are well-known, but we prefer to state them here for the reader's convenience. To formulate the following propositions, we return back to the initial numbering of shocks, denote the constant term in the active factors set of agents as w_0 and take $\rho_0 = 1$ and $B^0 = \alpha$. So, now i takes integer values from 0 to k. We will denote this set as I_0 and the corresponding set of subscripts taken from $I_0 = \{0, 1, ..., k\}$ as $\widetilde{I}_0$. Note that the constant term is always included as a factor in any active factors set, therefore 0 always belongs to I_0.

Proposition 5.1 *(Necessary and sufficient conditions for existence of a REE) Under Assumption 5.1, a REE solution exists if and only if agents' active factors set includes among others all w_i such that $B^i \neq 0$ in the reduced form and $rank(\rho_i A - I) = rank(\rho_i A - I, B^i)$ for i such that $\det(\rho_i A - I) = 0$ and $B^i \neq 0$.*

Proof. See Appendix C.1. □

Proposition 5.2 *(Necessary and sufficient conditions for existence and uniqueness of a REE): Under Assumption 5.1, a REE solution exists and is unique if and only if agents' active factors set includes, among others, all w_i such that $B^i \neq 0$ in the reduced form and for all w_i included, $\det(\rho_i A - I) \neq 0$.*

Proof. See Appendix C.1. □

So, in what follows we always assume that Assumption 5.1 and the necessary and sufficient conditions[9] for existence of a REE hold true. Basically, we assume that in both equations of their PLM, agents use at least all the regressors that appear in the right–hand side of the reduced form (5.4), and that the REE solution (either unique or

the discussion of which may be found, for example in Evans and Honkapohja [29]. In our terminology, for the situation of the RPE, one has to assume that agents do not include into their active factors set some of the factors that are present in a unique REE, that is, factors that correspond to non–zero coefficients in matrix B. Here we introduce the notion of the active factors set only to allow for considering the PLMs not only corresponding to the MSV, but also those PLMs that include more factors than enough to determine a unique REE. It is done to derive the "strong δ–stability" or "strong δ–instability result." (Compare to the notion of "strong E–stability" in the homogeneous learing literature.)

[9] The propositions above have a similar meaning to Proposition 1 of Honkapohja and Mitra [43]: again, the condition requires matrices participating in the derivation of the RE values of beliefs to be invertible. So, the above propositions stress that we are aware of cases when an REE may not exist and of the conditions that are required for its existence (and uniqness).

multiple) exists under this PLM. That is, in principle, we consider all possible PLMs that satisfy these conditions.

After specifying PLMs of agents and conditions for existence and uniqueness of the REE we are ready to introduce heterogeneous learning of agents in the economy considered and derive conditions for stability of the REE under this learning. Then we use these conditions to study stability under heterogenous learning in the general New Keynesian model when optimal monetary policy rules are applied.

5.3 Heterogeneous Learning and the Concept of δ–stability

The model (5.4) and (5.3) that we consider belongs to the class of multivariate forward–looking economic models. Thus we naturally employ the general framework and notation from Honkapohja and Mitra [43], who were the first to formulate the general criterion for stability of a multivariate forward–looking economy under heterogeneous learning.

Honkapohja and Mitra [43] consider the class of linear structurally heterogeneous forward-looking models with S types of agents with different forecasts presented by

$$y_t = \alpha + \sum_{h=1}^{S} A_h \hat{E}_t^h y_{t+1} + B w_t, \qquad (5.10)$$

$$w_t = F w_{t-1} + v_t, \qquad (5.11)$$

where y_t is an $n \times 1$ vector of endogenous variables, w_t is a $k \times 1$ vector of exogenous variables, v_t is white noise, $\hat{E}_t^h y_{t+1}$ are (in general, non-rational) expectations of the endogenous variable by agent type h, $M_w = \lim_{t\to\infty} w_t w_t'$ is positive definite, and F is such that w_t follows a stationary VAR process.

The PLM is presented by (5.6). Part of agent types, $h = \overline{1, S_0}$, is assumed to use the RLS learning algorithm, while the rest, $h = \overline{S_0 + 1, S}$, are assumed to use the SG learning algorithm.[10] Moreover, all of them are assumed to use possibly different degrees of responsiveness to the updating function that are presented by different degrees of inertia $\delta_i > 0$, constant coefficients before the common for all agents decreasing gain sequence in the learning algorithm.[11]

[10]Essentially, the part of agents using RLS are assumed to be more sophisticated in their learning, because from the econometric point of view, the RLS algorithm is more efficient since it uses information on the second moments.

[11]Honkapohja and Mitra [43] use a more general formulation of degrees of inertia.

It is worth noting that the model (5.4) and (5.3) that we consider belongs to the subclass of models considered by Honkapohja and Mitra [43], namely, a class of structurally homogeneous forward looking models. Structural heterogeneity in the setup of Honkapohja and Mitra [43] is expressed through matrices A_h, which are assumed to incorporate mass ζ_h of each agent type. That is, $A_h = \zeta_h \cdot \hat{A}_h$, where $\hat{A}_h$ is defined as describing how agents of type h respond to their forecasts. So these are the structural parameters characterizing a given economy. Those may be basic characteristics of agents, like the ones describing their preferences, endowments, and technology. Structural heterogeneity means that all $\hat{A}_h$'s are different for different types of agents. When $\hat{A}_h = A$ and $\sum \zeta_h = 1$, the economy is structurally homogenous.

When we apply conditions for a structurally homogeneous economy, $A_h = \zeta_h A$, where $\sum_{h=1}^{S} \zeta_h = 1$, and $1 > \zeta_h > 0$, to the model (5.10) and (5.11) considered by Honkapohja and Mitra [43], we get

$$y_t = \alpha + \sum_{h=1}^{S} A_h \hat{E}_t^h y_{t+1} + B w_t = \alpha + \sum_{h=1}^{S} \zeta_h A \hat{E}_t^h y_{t+1} + B w_t = \\ = \alpha + A \underbrace{\sum_{h=1}^{S} \zeta_h \hat{E}_t^i y_{t+1}}_{\hat{E}_t^{aver} y_{t+1}} + B w_t,$$

which is exactly the formulation of the structurally homogeneous model considered by Giannitsarou [37].[12] Thus conditions for stability valid for the (more general) class of structurally heterogeneous forward–looking models remain valid for the class of structurally homogeneous models.

After denoting $z_t = (1, w_t)$ and $\Phi_{h,t} = (a_{h,t}, \Gamma_{h,t})$, the formal presentation of the learning algorithms in this model can be written as follows.

RLS: for $h = \overline{1, S_0}$

$$\begin{aligned} \Phi_{h,t+1} &= \Phi_{h,t} + \alpha_{h,t+1} R_{h,t}^{-1} z_t \left(y_t - \Phi_{h,t}' z_t \right)' \\ R_{h,t+1} &= R_{h,t} + \alpha_{h,t+1} \left(z_{t-1} z_{t-1}' - R_{h,t} \right) \end{aligned} \qquad (5.12)$$

[12]Heterogeneous learning in the structurally homogeneous case was considered by Giannitsarou [37] for a more general class of self–referential linear stochastic models, which includes in itself the case of forward–looking models. Since our setup does not assume lagged endogenous variables, we concentrate on the structurally homogeneous case of forward–looking models that are a subclass of models considered by Giannitsarou [37] and at the same time are a special case of the setup of Honkapohja and Mitra [43].

SG: for $h = \overline{S_0 + 1, S}$

$$\Phi_{h,t+1} = \Phi_{h,t} + \alpha_{h,t+1} z_t \left(y_t - \Phi'_{h,t} z_t \right)' \tag{5.13}$$

Honkapohja and Mitra [43] show that stability of the REE, Φ_t, in this model is determined by stability of the ODE[13]:

$$\begin{aligned} \frac{d\Phi_h}{d\tau} &= \delta_h \left(T(\Phi')' - \Phi_h \right), h = \overline{1, S_0} \\ \frac{d\Phi_h}{d\tau} &= \delta_h M_z \left(T(\Phi')' - \Phi_h \right), h = \overline{S_0 + 1, S}, \end{aligned}$$

where $M_z = \lim_{t\to\infty} E z_t z_t'$.

The conditions for stability of this ODE give the general criterion for stability result for this class of models presented in Proposition 5 in Honkapohja and Mitra [43]. In the economy (5.10) and (5.11), the mixed RLS/SG learning (5.12) and (5.13) converges globally (almost surely) to the minimal state variable (MSV) solution if and only if matrices $D_1\Omega$ and $D_w\Omega_F$ have eigenvalues with negative real parts, where

$$\begin{aligned} D_1 &= \begin{pmatrix} \delta_1 I_n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \delta_S I_n \end{pmatrix}, \Omega = \begin{pmatrix} A_1 - I_n & \cdots & A_S \\ \vdots & \ddots & \vdots \\ A_1 & \cdots & A_S - I_n \end{pmatrix} \\ D_w &= \begin{pmatrix} D_{w1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{wS} \end{pmatrix}, \quad \begin{array}{l} D_{wh} = \delta_h I_{nk}, h = \overline{1, S_0} \\ D_{wh} = \delta_h \left(M_w \otimes I_n \right), h = \overline{S_0 + 1, S} \end{array} \\ \Omega_F &= \begin{pmatrix} F' \otimes A_1 - I_{nk} & \cdots & F' \otimes A_S \\ \vdots & \ddots & \vdots \\ F' \otimes A_1 & \cdots & F' \otimes A_S - I_{nk} \end{pmatrix}, \end{aligned} \tag{5.14}$$

with $\otimes$ denoting the Kronecker product.

Note, that agents in the setup of Honkapohja and Mitra [43] are assumed to use PLMs that correspond to the so-called MSV solution, i.e., include all factors that appear in the right hand side of the reduced form. However, Honkapohja and Mitra [43] in their proof of conditions for stability of the system do not have restrictions on the matrix B.

[13] In the general case, to obtain the associated ODE, one has to take the math expectation of the RHS term (at the gain sequence) from the stochastic recursive algorithm (SRA) specification of a learning algorithm, with respect to the limiting distribution of the state vector. See Ch. 6.2 in Evans and Honkapohja [29] for assumptions on the learning rule and state dynamics that have to hold so that we are able to apply the theory on SRA and local convergence analysis and the general formula for ODE (6.5) on p. 126.

This means that we may, in principle, consider additional factors in learning that enter the reduced form with zero coefficients in matrix B for all agents. This means that we may consider the criterion conditions for all possible PLMs that include (among others) all factors that appear in the right hand side of the reduced form, satisfying conditions for existence specified in the previous chapter.

Kolyuzhnov [51] shows that in the "diagonal" environment, namely

$$F = diag(\rho_1, ..., \rho_k), M_w = diag\left(\frac{\sigma_1^2}{1-\rho_1^2}, ..., \frac{\sigma_k^2}{1-\rho_k^2}\right), \tag{5.15}$$

which we consider in this paper, the problem of finding stability conditions for both $D_1\Omega$ and $D_w\Omega_F$ is simplified to finding stability conditions for $D_1\Omega$ and $D_1\Omega_{\rho l}$,where Ω_{ρ_l} is obtained from Ω by substituting all A_h with $\rho_l A_h$, where $|\rho_l| < 1$ as w_t follows a stationary VAR(1) process.

$$\Omega_{\rho_l} = \begin{pmatrix} \rho_l A_1 - I_n & \cdots & \rho_l A_S \\ \vdots & \ddots & \vdots \\ \rho_l A_1 & \cdots & \rho_l A_S - I_n \end{pmatrix}, \forall l = 0, ..., k, (\rho_0 = 1). \tag{5.16}$$

Kolyuzhnov [51] uses a special blocked—diagonal structure of matrix D_1, which is the feature of the dynamic environment in this class of models. In a sense, these positive diagonal D–matrices may now be called positive blocked—diagonal δ–matrices. This makes it possible to formulate the concept of δ–stability by analogy to the terminology of the concept of D–stability about matrices that remain stable under multiplication by a diagonal matrix with positive elements, studied for example in Johnson [44].

Definition 5.2 *Given n, the number of endogenous variables, and S, the number of agent types, δ–stability is defined as stability of the economy under structurally heterogeneous mixed RLS/SG learning for any (possibly different) degrees of inertia of agents, $\delta > 0$.*

δ–stability, thus formulated, has the same meaning in models with heterogeneous learning described above as has the $E-$stability condition in models with homogeneous RLS learning. The $E-$stability condition is a condition for asymptotic stability of an REE under homogeneous RLS learning. The REE of the model is stable if it is locally asymptotically stable under the following ODE:

$$\frac{d\theta}{d\tau} = T(\theta) - \theta,$$

where θ are the estimated parameters from agents PLMs, $T(\theta)$ is a mapping of the PLM parameters into the parameters of the actual law of motion (ALM), which is obtained when we plug the forecast functions based on the agents' PLMs into the reduced form of the model, and τ is a "notional" ("artificial") time. The fixed point of this ODE is the REE of the model.[14]

Note that the δ–stability concept comprises stability under the three types of heterogeneous learning considered by Giannitsarou [37]. It is worth noting that in the case of heterogeneous learning in a structurally homogeneous economy, which we employ in the current setup, the criterion of Honkapohja and Mitra [43] is simplified to conditions on the Jacobians considered by Giannitsarou [37]. First, as has been discussed before, to get the structurally homogeneous economy as discussed before, one has to replace A_i in the setup of Honkapohja and Mitra [43] with $\zeta_i A$. After that, one has to make the following simplifications in the setup corresponding to a particular type of heterogeneous learning considered.

The first type of heterogeneous learning is characterized by different initial perceptions of agents and equal degrees of inertia. This type is termed transiently heterogeneous learning by Honkapohja and Mitra [43]. The condition for stability under this learning is easily derived from the criterion above by setting all δ's to be equal, and setting S_0 to S or to 0 in order to get transiently heterogeneous RLS or SG learning, respectively.

The second type of heterogeneous learning considered by Giannitsarou [37] is such that agents use different degrees of inertia and the same type of learning algorithm (RLS or SG). This is what Honkapohja and Mitra [43] call persistently heterogeneous learning in a weak form. The Jacobians for this case are easily derived by setting S_0 to S or to 0 in order to get heterogeneous RLS or SG learning, respectively, and allowing for possibly different δ's.

The third type of heterogeneous learning considered by Giannitsarou [37] is characterized by possibly different initial perceptions, possibly different degrees of inertia, and by different agents using different learning algorithms (RLS or SG). Such kind of learning

[14]Notice that δ–stability conditions on the Jacobian in the general forward–looking model of Honkapohja and Mitra [43] do not depend on the particular equilibrium point (in case of multiple equilibria), because the system of differential equations is linear in this setup, in which case the first derivatives of the RHS of the associated ODE do not depend on a particular value of a RE equilibrium. So if stability conditions are satisfied for a given Jacobian, then all equilibrium points are stable. Convergence to a particular point depends on the initial conditions. In this paper we do not consider how equilibrium selection is made.

Type of heterogeneity	Type of learning	Assumptions in the general H&M (2006) model	
		structurally heterogeneous	structurally homogeneous
		$A_h = \zeta_h \hat{A}_h$	$A_h = \zeta_h A$
I Different initial perceptions (transiently heterogeneous learning)	RLS SG	$\delta_h = \delta$ for all h, $S_0 = S$ $\delta_h = \delta$ for all h, $S_0 = 0$	
II Different degrees of inertia (persistently heterogeneous learning in a weak form)	RLS SG	$S_0 = S$ $S_0 = 0$	
III Different learning algorithms (persistently heterogeneous learning in a strong form)	RLS and SG		

Table 5.1: Types of heterogeneity in learning.

Honkapohja and Mitra [43] call persistently heterogeneous learning in a strong form. The stability Jacobians for this case are derived by writing the general criterion for stability for the structurally homogeneous case, i.e., by setting $A_i = \zeta_i A$.

The relation between the above–described formulations of the types of heterogeneous learning by Giannitsarou [37] and by Honkapohja and Mitra [43] can be conveniently summarized in the following table[15]:

Notice that in the "diagonal" case (5.15), δ–stability does not depend on S_0. Thus if the economy (5.10), (5.11) and (5.15) is δ–stable, it is stable under all three types of heterogeneous learning and under RLS and SG homogeneous learning.

[15]Note that there is one type of heterogeneous learning that was not introduced by Giannitsarou [37] and is introduced here. It is heterogeneity in degrees of inertia under which all types of agents use the SG learning algorithm. Although Honkapohja and Mitra [43] have the general criterion for stability in this case (as discussed above), their formulation includes only forward–looking models. In the general setup of self-referential structurally homogeneous models of Giannitsarou [37], the stability conditions under such type of learning (in Giannitsarou [37] notation, naturally extended from her proofs) would depend on the stability of matrix $J_2^{SG}(\Phi_f) = diag(\delta_1, ..., \delta_S) \otimes I \otimes M(\Phi_f) J_1^{LS}(\Phi_f)$, where Φ_f is an REE, $M(\Phi_f)$ is defined similarly to M_z and $J_1^{LS}(\Phi_f)$ is a Jacobian that defines stability in case of the first type of heterogeneity (different initial perceptions of agents) when all agents use RLS learning. For details, see Giannitsarou [37] . Again, it is clear that in the forward–looking case these conditions for stability fall under the general stability criterion of Honkapohja and Mitra [43] with $S_0 = 0$ (see the table above).

5.4 Conditions for δ–stability of Structurally Homogeneous Models

After establishing the universal role of the concept of δ–stability for stability under all three types of heterogeneous learning discussed above, we present necessary and sufficient conditions. First, we provide the reader with a set of sufficient conditions for δ–stability applicable to our setup, that is, for a class of structurally homogeneous models. We present (without proofs) the so–called aggregate economy sufficient condition for the case of a structurally homogeneous model and the "same sign" sufficient condition for the case of a structurally heterogeneous bivariate economy that were derived in Kolyuzhnov [51]

Proposition 5.3 *For the structurally homogeneous economy (5.4) and (5.3) to be δ–stable, it is sufficient that at least one of the following maximal aggregated β–coefficients (which are the coefficients before the expectation term of a one–dimensional forward–looking aggregate economy model. For details see Kolyuzhnov [51]):* $\max_i \sum_j |a_{ij}|$ *and* $\max_j \sum_i |a_{ij}|$ *are less than one, where a_{ij} denotes an element in the i^{th} row and the j^{th} column of A.*

Proposition 5.4 *In case $n = 2$, the economy (5.10), (5.11) and (5.15) is δ–stable if the corresponding matrix Ω, defined in (5.14), is stable and the following "same sign" condition holds true:*

$$\det(-\rho_l A_i) \geq 0, [\det mix(-\rho_l A_i, -\rho_l A_j) + \det mix(-\rho_l A_j, -\rho_l A_i)] \geq 0, i \neq j, M_1(-\rho_l A_i) \geq 0$$

or

$$\det(-\rho_l A_i) \leq 0, [\det mix(-\rho_l A_i, -\rho_l A_j) + \det mix(-\rho_l A_j, -\rho_l A_i)] \leq 0, i \neq j, M_1(-\rho_l A_i) \leq 0,$$

$$\forall l = 0, 1, ..., k, (\rho_0 = 1),$$

where $mix(-\rho_l A_i, -\rho_l A_j)$ denotes a matrix of structural parameters of a pairwise–mixed economy and is composed by mixing columns of a pair of matrices $\rho_l A_i$, $\rho_l A_j$, for any i, $j = \overline{1, S}$.

It is also possible to derive some necessary conditions and sufficient conditions of δ–stability in the structurally homogeneous case in terms of the values of eigenvalues of the matrix of structural parameters of the reduced form, A. It is possible by the direct application of the characteristic equation approach, when one requires that all the roots of the polynomial (that are eigenvalues of the Jacobian matrix) be less than zero for stability, the latter being equivalent to the well-known Routh–Hurwitz conditions.

Proposition 5.5 *If all eigenvalues of A are real and less than one, then the structurally homogeneous system (5.4) and (5.3) with two agents is δ–stable, that is, stable under the three types of heterogeneous learning: agents with different initial perceptions with RLS or SG learning, agents with possibly different degrees of inertia with RLS or SG learning, and agents with different learning algorithms, RLS and SG. For the structurally homogeneous system (5.4) and (5.3) with any number of agents to be $\delta-$stable, it is necessary that all real roots of A be less than one. This gives a test for non–δ–stability.*

Proof. See Appendix C.1. □

In the proof of the proposition above, using the structure of the Jacobian matrices in our setup, we have derived a sufficient condition for stability under all three types of heterogeneous learning with two agent types. We did this using the criterion for stability of Honkapohja and Mitra [43]. For the case of real roots of A, we have shown that in this setup, the analysis of stability of a particular Jacobian turns into the analysis of stability of A, which gives us very simple eigenvalues conditions. Also, using the general criterion of Honkapohja and Mitra [43], we have proved here the necessary conditions for δ–stability (the failure of which is sufficient for non–δ–stability) for the case of an arbitrary number of agent types.

5.5 Optimal Policy Rules and the Structure of the Reduced Forms

After deriving and stating the conditions for stability under the three types of heterogeneous learning discussed in the previous section, we are ready to study the general New Keynesian model (5.1), (5.2) and (5.3) for stability under heterogenous learning when optimal monetary policy rules are applied. Here we describe the types of optimal policy rules that are analyzed in this study.

The policy maker is assumed to use the loss function minimization problem, which comes from the flexible inflation targeting approach (a policy regime adopted in several countries in the 1990s), described and defended by Svensson [62]. The central bank here has two options: adopt a discretionary policy, by solving the problem every period, or commit to a rule which is once and for all derived from the minimization of the infinite horizon loss function. Svensson [62] and Cecchetti [14] advocate the first option, which is essentially commitment to a certain behavior (minimizing the loss function) with reconsidering the optimal rule every period, so that to take into account new information. They provide various arguments, like inefficiency (in general) of instrument rules designed to respond only to target variables or the way monetary policy decisions are made in practice.

The infinite horizon loss function of the policy maker for the flexible inflation targeting approach looks as follows.

$$\frac{1}{2}E_t\sum_{i=0}^{\infty}\beta^i\left[\alpha\left(x_{t+i}-\bar{x}\right)^2+\left(\pi_{t+i}-\bar{\pi}\right)^2\right]$$

According to the discussion above, we assume the discretionary policy of the policy maker and the problem of minimizing the loss function simplifies to solving each period

$$\min\frac{1}{2}\left[\alpha\left(x_t-\bar{x}\right)^2+\left(\pi_t-\bar{\pi}\right)^2\right]+R_t \tag{5.17}$$

subject to

$$\pi_t = c_2+\lambda x_t+F_t$$

(the central bank takes the remainder terms of the loss function R_t, and the constraint $F_t=\beta\widehat{E}_t\pi_{t+1}+\chi_2 w_t$ as given).

The classification below of the loss–function–optimization–based rules into fundamentals–based and expectations–based rules provided below is due to Evans and Honkapohja [31]. The derivation of these rules and the corresponding reduced forms is done by Evans and Honkapohja [31] for a slightly more narrow setup than is assumed here (we assume general structure of autoregressive shocks), therefore in the derivations that follow below we basically repeat their steps extending them for our setup.

5.5.1 Expectations–based Optimal Policy Rules

The expectations–based policy rule implies the central bank's reaction to (possibly non-rational) expectations of private agents, assuming that these expectations are observable (or can be estimated). Its general form is $i_t = \delta_0 + \delta_\pi \hat{E}_t \pi_{t+1} + \delta_x \hat{E}_t x_{t+1} + \delta_w' w_t$. The coefficients of this rule are obtained by solving the equilibrium conditions: structural equations with non–rational expectations of private agents (5.1) and (5.2) and the first order conditions (FOC) of the optimization problem of the central bank (5.17), $\lambda(\pi_t - \bar{\pi}) + \alpha(x_t - \bar{x}) = 0$. Thus, the expectations–based policy rule looks as follows:

$$\begin{aligned} & i_t = \delta_0 + \delta_\pi \hat{E}_t \pi_{t+1} + \delta_x \hat{E}_t x_{t+1} + \delta_w' w_t, \text{ where} \\ & \delta_0 = -\left(\lambda^2 + \alpha\right)^{-1} \phi^{-1} \left(\lambda\bar{\pi} + \alpha\bar{x} - \lambda c_2 - \left(\alpha + \lambda^2\right) c_1\right), \\ & \delta_\pi = 1 + \left(\lambda^2 + \alpha\right)^{-1} \phi^{-1} \lambda\beta,\ \delta_x = \phi^{-1},\ \delta_w = \phi^{-1}\chi_1 + \left(\lambda^2 + \alpha\right)^{-1} \phi^{-1} \lambda \chi_2 \end{aligned} \tag{5.18}$$

After plugging this policy rule into the IS curve equation, we get the following reduced form.

$$\begin{aligned} y_t &= c^E + A^E \hat{E}_t y_{t+1} + \chi^E w_t, \\ w_t &= F w_{t-1} + \nu_t, \\ y_t &= \begin{bmatrix} \pi_t & x_t \end{bmatrix}', \text{ where } F = diag(\rho_i),\ |\rho_i| < 1, \nu_{i_t} \sim iid\left(0, \sigma_i^2\right), i = \overline{1, n}, \\ A^E &= \begin{pmatrix} \beta\alpha\left(\lambda^2 + \alpha\right)^{-1} & 0 \\ -\beta\lambda\left(\lambda^2 + \alpha\right)^{-1} & 0 \end{pmatrix}, \\ c^E &= \begin{pmatrix} c_2 + \lambda\left(c_1 - \phi\delta_0\right) \\ c_1 - \phi\delta_0 \end{pmatrix}, \chi^E = \begin{pmatrix} \chi_2' \left[1 - \frac{\lambda^2}{\lambda^2 + \alpha}\right] \\ -\frac{\lambda^2}{\lambda^2 + \alpha} \chi_2' \end{pmatrix} \end{aligned} \tag{5.19}$$

Note that the REE solution is not needed either for deriving matrix A^E, or for deriving the coefficients of the optimal expectations–based policy rule. The REE solution will be needed for deriving the optimal fundamentals–based policy rule, and therefore will be derived in the corresponding part of the text.

5.5.2 Fundamentals–based Optimal Policy Rules

In general, the fundamentals–based policy rule (not necessarily optimal) has the form

$$i_t = \psi_0 + \sum_{i=1}^{n} \psi_{w_i} w_{i_t} = \psi_0 + \psi_w' w_t \tag{5.20}$$

Later we show that there exists a unique fundamentals–based optimal policy rule in this setup and derive this rule.

Plugging this policy rule into the structural form (5.1) and (5.2), we get the following reduced form:

$$
\begin{aligned}
y_t &= c^F + A^F \widehat{E}_t y_{t+1} + \chi^F w_t, \\
w_t &= F w_{t-1} + \nu_t, \\
y_t &= \begin{bmatrix} \pi_t & x_t \end{bmatrix}', \text{ where } F = diag(\rho_i),\ |\rho_i| < 1, \nu_{i_t} \sim iid\left(0, \sigma_i^2\right), i = \overline{1,n}, \\
A^F &= \begin{pmatrix} \beta + \lambda\phi & \lambda \\ \phi & 1 \end{pmatrix}, \qquad (5.21) \\
c^F &= \begin{pmatrix} c_1 - \phi\psi_0 \\ c_2 + \lambda\left(c_1 - \phi\psi_0\right) \end{pmatrix}, \ \chi_F = \begin{pmatrix} \lambda\left(-\phi\psi_w' + \chi_1'\right) + \chi_2' \\ -\phi\psi_w' + \chi_1' \end{pmatrix}.
\end{aligned}
$$

The optimal fundamentals–based rule, under the central banks' discretionary policy, is obtained from the loss function minimization, with the central bank assuming that private agents have RE. With the REE structure being $y_t = a + \Gamma w_t$, its general form is $i_t = \psi_0 + \psi_w' w_t$, where w_t is a vector of exogenous variables. Using the equilibrium conditions (economy's structural equations (5.1) and (5.2), with the REE structure entering them and the FOC of the central bank's optimization problem), we obtain the coefficients of the REE and of the optimal fundamentals–based policy rule.

To get the REE, one has to write the ALM using the Phillips curve (5.2), the FOC of the central bank's optimization problem and the PLM in the general form, $y_t = a + \Gamma w_t$, and then according to the RE principle, equate coefficients of the resulting ALM (T–mapping) with the corresponding coefficients of the PLM. The resulting ALM looks like

$$
\begin{aligned}
\pi_t &= \frac{c_2 + \lambda\left[\lambda\bar{\pi} + \alpha\bar{x}\right]}{\lambda^2 + \alpha} + \frac{\alpha\beta}{\lambda^2 + \alpha}\left[a_1 + \gamma_{11}\rho_1 w_{1t} + ... + \gamma_{1n}\rho_n w_{nt}\right] + \frac{\alpha}{\lambda^2 + \alpha}\chi_2' w_t \\
x_t &= \frac{\lambda\bar{\pi} + \alpha\bar{x}}{\alpha} - \frac{\lambda}{\alpha}\pi_t
\end{aligned}
$$

and the REE looks like

$$
\begin{aligned}
\pi_t &= a_1^* + \sum_{i=1}^{n} \gamma_{1i}^* w_{it} \qquad (5.22) \\
x_t &= a_2^* + \sum_{i=1}^{n} \gamma_{2i}^* w_{it}, \text{ where} \\
a_1^* &= \frac{c_2 + \lambda\left[\lambda\bar{\pi} + \alpha\bar{x}\right]}{\lambda^2 + \alpha\left(1 - \beta\right)}, a_2^* = \frac{\lambda\bar{\pi} + \alpha\bar{x}}{\alpha} - \frac{\lambda}{\alpha}a_1^* = \frac{-\frac{\lambda}{\alpha}c_2 + \left(1 - \beta\right)\left[\lambda\bar{\pi} + \alpha\bar{x}\right]}{\lambda^2 + \alpha\left(1 - \beta\right)}, \\
\gamma_{1i}^* &= \frac{\alpha\chi_{2i}\rho_i}{\alpha\left(1 - \beta\rho_i\right) + \lambda^2}, \gamma_{2i}^* = -\frac{\lambda}{\alpha}\gamma_{1i}^* = -\frac{\lambda\chi_{2i}\rho_i}{\alpha\left(1 - \beta\rho_i\right) + \lambda^2}, i = \overline{1,n}.
\end{aligned}
$$

To get the *optimal* fundamentals–based policy rule, one has to express i_t using the IS curve (5.1), plugging in it the REE solution (5.22) derived above.

$$i_t = -\frac{1}{\phi}\left(a_2^* + \sum_{i=1}^{n} \gamma_{2i}^* w_{it}\right) + \left(a_1^* + \sum_{i=1}^{n} \gamma_{1i}^* \rho_i w_{it}\right) + \frac{1}{\phi}\left(a_2^* + \sum_{i=1}^{n} \gamma_{2i}^* \rho_i w_{it}\right) + \frac{1}{\phi}\chi_1' w_t$$

As a result, the optimal fundamentals–based policy rule looks like

$$\begin{aligned} i_t &= \psi_0^* + \psi_w^{*\prime} w_t, \text{ where} \qquad (5.23) \\ \psi_0^* &= a_1^*,\ \psi_w^* = \frac{1}{\phi}\left[\begin{pmatrix} \gamma_{21}(\rho_1 - 1) & \dots & \gamma_{2n}(\rho_n - 1) \end{pmatrix} + \chi_1\right] + \begin{pmatrix} \gamma_{11}\rho_1 & \dots & \gamma_{1n}\rho_n \end{pmatrix}. \end{aligned}$$

In both cases of optimal monetary policy rules, we plug the corresponding policy rule into the structural equations and obtain the corresponding reduced form of the model. These reduced forms were studied for stability under homogeneous RLS learning in the Clarida, Gali, and Gertler [19, 20] formulation of the New Keynesian model by Evans and Honkapohja [31] , who derived the stability results for the expectations–based rule and the instability results for the fundamentals–based rule. We study stability and instability for the two categories of rules under the heterogeneous learning of private agents in the general setup of the New Keynesian model (5.1), (5.2) and (5.3).

5.6 Stability Problem in the New Keynesian Model

After deriving the reduced forms corresponding to the optimal monetary policy rules, we are ready to check them for δ–stability. To do this we have to test the resulting matrix A of the reduced form (5.19) or (5.21) for the applicability of the sufficient and necessary conditions for δ–stability. For the situation of the optimal expectations–based policy rule we have the following result.

Proposition 5.6 *The general New Keynesian model with a stationary AR(1) observable shocks process (5.1), (5.2) and (5.3) is δ–stable when the optimal expectations–based policy rule (5.18) is applied.*[16]

Proof. We know that the corresponding A matrix in the optimal expectations–based policy rule case is $A^E = \begin{pmatrix} \beta\alpha\left(\lambda^2+\alpha\right)^{-1} & 0 \\ -\beta\lambda\left(\lambda^2+\alpha\right)^{-1} & 0 \end{pmatrix}$. Using the sufficient condition in

[16]This result is not very surprising as Evans, Honkapohja, and Williams [32] have a convergence result under the optimal expectations–based policy rule when all agents use SG learning.

Proposition 5.4, we have that Ω is stable, since its eigenvalues are determined from the following characteristic equation $\det\left(A^E - I_2(1+\mu)\right)(1+\mu)^{2(S-1)} = 0$ and therefore, are equal to -1 and $\beta\alpha\left(\lambda^2+\alpha\right)^{-1} - 1$, i. e., are negative, and we have that $\det(-\rho_l A_i) = 0$, $[\det mix(-\rho_l A_i, -\rho_l A_j) + \det mix(-\rho_l A_j, -\rho_l A_i)] = 0, i \neq j, M_1(-\rho_l A_i) = -\rho_l \zeta_h \beta\alpha\left(\lambda^2+\alpha\right)^{-1} \geq (\leq) 0$,for all $l = 0, 1, ..., k$ $(\rho_0 = 1)$, so the "same sign" condition holds true. Notice that using the "aggregate economy" sufficient condition from Proposition 5.3, we can write two aggregate β–coefficients in the expectations–based policy rule case. These are $\beta_1^{\max} = \max_i \sum_j |a_{ij}| = \max\left\{\beta\alpha\left(\lambda^2+\alpha\right)^{-1}, \beta\lambda\left(\lambda^2+\alpha\right)^{-1}\right\}$ and $\beta_2^{\max} = \max_j \sum_i |a_{ij}| = \beta(\alpha+\lambda)\left(\lambda^2+\alpha\right)^{-1}$. It is clear that both coefficients are less than one if $\lambda \geq 1$. So, the "aggregate economy" sufficient condition for δ–stability is a more restrictive condition compared to the "same sign" condition since it requires additional assumptions on the structure of the economy. However, it can be with success applied in more than two dimensional economies where similar "same sign" conditions are not sufficient for δ–stability (see Kolyuzhnov [51]). □

Note that Evans and Honkapohja [31] have a similar result for homogeneous learning. The proposition below presents the instability result for the situation of the fundamentals–based monetary policy rule.

Proposition 5.7 *The general New Keynesian model with a stationary AR(1) observable shocks process (5.1), (5.2) and (5.3) is non–δ–stable when the fundamentals–based policy rule (5.20) (as well as the optimal fundamentals–based policy rule (5.23)) is applied.*

Proof. We know that the corresponding matrix A in the fundamentals–based policy rule case is $A^F = \begin{pmatrix} \beta+\lambda\phi & \lambda \\ \phi & 1 \end{pmatrix}$. Using the "eigenvalues" necessary condition from Proposition 5.5,[17] we get the eigenvalues of this matrix:.$\mu_{1,2} = 1 + \frac{\beta+\lambda\phi-1}{2} \pm \sqrt{\left(\frac{\beta+\lambda\phi-1}{2}\right)^2 + \lambda\phi}$. Both of these eigenvalues are real and eigenvalue $\mu_1 = 1 + \frac{\beta+\lambda\phi-1}{2} + \sqrt{\left(\frac{\beta+\lambda\phi-1}{2}\right)^2 + \lambda\phi}$ is greater than one. So, the sufficient condition for non–δ–stability is satisfied. □

Again, Evans and Honkapohja [31] have a similar result for homogeneous learning.

[17] In principle, we could also use our necessary conditions for δ–stability (derived in Kolyuzhnov [51]) to show the instability of the fundamentals–based rule. However, these may be more difficult to check than the necessary conditions on eigenvalues that we derived in this paper. Besides, our eigenvalues necessary conditions work for the case of an arbitrary number of agent types.

Proposition 5.6 means that the REE in this model, resulting after implementing the optimal expectations–based policy rule, is stable under the recursive least squares and the stochastic gradient homogeneous learning and the three types of heterogeneous learning: agents with different initial perceptions with the RLS or SG learning, agents with different degrees of inertia with RLS or SG learning, and agents with different learning algorithms, RLS and SG. Proposition 5.7 claims that the REE of this model with the fundamentals–based policy rule is always unstable under any type of heterogeneous and homogeneous learning of agents.

5.7 Conclusion

We have used the environment of the New Keynesian model to explore the question of stability of two categories of optimal monetary policy rules under the assumption of heterogeneous learning of private agents.

These two categories were introduced by Evans and Honkapohja [31], and this division is based on the assumption about the central bank's perception of private agents' expectations: RE or possibly non-rational. Under the central bank assuming private agents to have RE, the fundamentals–based rule is obtained, while the case of the central bank assuming possibly non-rational expectations of private agents results in the fundamentals–based rule.

The purpose of this research was, on the one hand, to explore whether, given structural homogeneity of the model, heterogeneity in learning of agents influences the stability results implied by the application of either of the two categories of policy rules.

Using the general criterion for stability of Honkapohja and Mitra [43] and the sufficient δ–stability conditions derived in Kolyuzhnov [51] for the case of heterogeneous learning, we obtain results similar to those obtained by Evans and Honkapohja [31] for the case of homogeneous learning. In particular, under the fundamentals–based policy rule, the model economy is always unstable, so there is no convergence to the associated REE of the model, while there is stability under the optimal expectations–based rule and the economy converges to the REE corresponding to the optimal monetary policy without commitment.

The above–described results have been obtained using only the structure of the model, so there is no dependence on heterogeneity of any type considered. This implies that in the New Keynesian model, the stability results are independent of heterogeneity in learning, so the representative agent hypothesis is applicable in this setup.

The method of analysis presented in this paper allows us to check the applicability of this hypothesis in the case of heterogeneous leaning of private agents in the New Keynesian economy under Taylor–type rules (the case of internal central bank forecasting), which do not fall under the classification of Evans and Honkapohja [31]. This issue will be considered in a separate study.

Chapter 6

Afterword

My thesis makes a contribution to the economic literature on adaptive learning, in particular, to the areas of escape dynamics and heterogeneous learning. In the field of escape dynamics, I (together with Anna Bogomolova and Sergey Slobodyan) have developed a new way of calculating escape dynamics characteristics using the continuous–time approximation of the original discrete–time dynamics of the model, thus resolving the theoretical and computational problems associated with the discrete–time approach considered in CWS. The developed approach is presented in the first chapter of the thesis. The second chapter compares the behavior of the RLS and SG algorithms with constant gain in terms of dynamics (namely, mean and escape dynamics) under learning around the point of SCE and shows that the behavior of these learning algorithms substantially differs in terms of escape dynamics.

The third and fourth chapters make a contribution to the area of heterogeneous learning. In the third chapter I have derived conditions for stability of a structurally heterogeneous economy under heterogeneous learning in the form of mixed RLS/SG learning with (possibly) different degrees of inertia of agents. These conditions have strong theoretical and practical implementation. In terms of the theory, it is shown that these conditions are formulated using such theoretically reasonable concepts like an aggregate economy and a subeconomy and relate the concept of stability under homogeneous learning (E–stability) to the concept of stability under heterogeneous learning (δ–stability). From the practical point of view, it is shown, on an example of two types of OLG models, that it is very easy to test an economy for stability under heterogeneous learning using the conditions derived in this chapter.

In the fourth chapter, using the results on δ–stability derived in the third paper, I (together with Anna Bogomolova) show that the fundamentals based monetary policy rule is unstable and optimal expectations based monetary policy rule is stable in the general setup of the New Keynesian model under any type of heterogeneous private agents learning considered by Giannitsarou [37], thus extending the result of Evans and Honkapohja [31] derived for the case of representative private agent (homogeneous) learning.

The directions for future contributions are promising and follow directly from the previous contribution. One of the possible directions in the area of escape dynamics is to consider the case of mixed RLS/SG learning with constant gain in the Phelps problem considered in the first chapter in order to see what the resulting dynamics is[1]. Another direction in this area is to apply the developed continuous–time approach to another model. In this sense, one has to look for economically sensible models with better averaging (compared to the Phelps problem) for economically plausible gain values in order to apply the large deviations theory characteristics of escape time.

In particular, the large deviations theory analysis can be applied in models where one may observe new data and learn almost continuously due to an almost continuous data flow. Such a situation is typical for financial markets and currency exchange rates. For example, the continuous–time approach can be applied to the model of Aghion, Baccheta and Banerjee [1] of currency crises with the government and the public sector learning that was introduced into this model by Cho and Kasa [15], in which it is not possible to analytically characterize escape dynamics using the discrete–time approach due to its computational intensity. Cho and Kasa [15] used the continuous–time approximation derived in the first chapter of this thesis to qualitatively explain the dynamics in their model with two-sided learning. It could be also of interest to apply the continuous–time approach to the extension of the Sargent [61] model, i.e., to the case of the Dynamic Phillips Curve.

In the area of heterogenous learning of agents, the direction of my nearest future research (besides other directions that include stability of cycles and sunspot equilibria under heterogeneous learning) is to apply the derived conditions to study Taylor rules for stability under heterogeneous learning. Moreover, it is possible to consider a more general case of the internal central bank forecasting (that includes Taylor rules) in a situation

[1] The author thanks Kaushik Mitra for suggesting this direction of research.

of heterogeneous learning of private agents that develops the model of Honkapohja and Mitra [42], as Honkapohja and Mitra [42] consider only the situation of a representative private agent.

Bibliography

[1] Phillipe Aghion, Phillipe Bacchetta, and Abhijit Banerjee. Currency crises and monetary policy in an economy with credit constraints. *European Economic Review*, 45:1121–1150, 2001.

[2] Jasmine Arifovic. Genetic algorithm learning and the cobweb model. *Journal of Economic Dynamics and Control*, 18:3–28, 1994.

[3] R. J. Barro and D. B. Gordon. A positive theory of monetary policy in a natural-rate model. *Journal of Political Economy*, 91:589–610, 1983.

[4] R. J. Barro and D. B. Gordon. Rules, discretion, and reputation in a model of monetary policy. *Journal of Monetary Economics*, 12:101–121, 1983.

[5] Ken Binmore and Larry Samuelson. Muddling through: Noisy equilibrium selection. *Journal of Economic Theory*, 74:235–265, 1997.

[6] Anna Bogomolova and Dmitri Kolyuzhnov. Optimal monetary policy rules: The problem of stability under heterogeneous learning. Manuscript, CERGE-EI, 2006.

[7] Andrei N Borodin and Paavo Salminen. *Handbook of Brownian Motion — Facts and Formulae*. Birkhauser Verlag, Basel Boston Berlin, 1996.

[8] Stephen Boyd. Lecture notes for EE263, 2004. Stanford lecture notes.

[9] Margaret M Bray. Learning , estimation, and the stability of rational expectations equilibria. *Journal of Economic Theory*, 26:318–339, 1982.

[10] Margaret M Bray and Nathan E Savin. Rational expectations equilibria, learning, and model specification. *Econometrica*, 54(5):1129–60, 1986.

[11] James Bullard and Kaushik Mitra. Learning about monetary policy rules. *Journal of Monetary Economics*, 49(6):1105–29, 2002.

[12] James B Bullard and In-Koo Cho. Escapist policy rules. Forthcoming, Journal of Economic Dynamics and Control, 2005.

[13] G. A. Calvo. Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics*, 12:383–398, 1983.

[14] Stephen G. Cecchetti. Making monetary policy: Objectives and rules. *Oxford Review of Economic Policy*, 16(4):43–59, 2001.

[15] In-Koo Cho and Kenneth Kasa. Learning dynamics and endogenous currency crises. *Macroeconomic Dynamics*, 12:257–285, 2008.

[16] In-Koo Cho and Thomas J. Sargent. Neural networks for encoding and adapting in dynamic economies. In H. M. Amman, D. A. Kendrick, and J. Rust, editors, *Handbook of Computational Economics*, pages 441–470. 1996.

[17] In-Koo Cho, Noah Williams, and Thomas J Sargent. Escaping nash inflation. *Review of Economic Studies*, 69(1):1–40, 2002.

[18] L. J. Christiano, M Eichenbaum, and C Evans. Nominal rigidities and the dynamic effects of a shock to monetary policy. *NBER Working Paper Series*, w. p. 8403, 2001.

[19] Richard Clarida, Jordi Gali, and Mark Gertler. The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37:1661–1707, 1999.

[20] Richard Clarida, Jordi Gali, and Mark Gertler. A simple framework for international monetary policy analysis. *Journal of Monetary Economics*, 49:879–904, 2002.

[21] Mohammed Dahleh, Munther A Dahleh, and George Verghese. Lectures on dynamic systems and control. MIT lecture notes, 2004.

[22] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. Springer, New York Berlin Heidelberg, 1998.

[23] Paul Dupuis and Harold Kushner. Stochastic approximations via large deviations: Asymptotic properties. *SIAM Journal of Control and Optimization*, 23:675–696, 1985.

[24] Paul Dupuis and Harold J Kushner. Stochastic approximation and large deviations: Upper bounds and w.p.1 convergence. *SIAM Journal on Control and Optimization*, 27:1108–1135, 1989.

[25] George W Evans and Seppo Honkapohja. Learning, convergence, and stability with multiple rational expectations equilibria. *European Economic Review*, 38:1071–1098, 1994.

[26] George W Evans and Seppo Honkapohja. On the local stability of sunspot equilibria under adaptive learning rules. *Journal of Economic Theory*, 64:142–161, 1994.

[27] George W Evans and Seppo Honkapohja. Local convergence of recursive learning to steady states and cycles in stochastic nonlinear models. *Econometrica*, 63:195–206, 1995.

[28] George W. Evans and Seppo Honkapohja. Learning dynamics. In John Tailor and M. Woodford, editors, *Handbook of Macroeconomics*, pages 449–542. North Holland, Amsterdam, 1999.

[29] George W. Evans and Seppo Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, NJ, 2001.

[30] George W Evans and Seppo Honkapohja. Adaptive learning and monetary policy design. *Journal of Money, Credit, and Banking*, 35(6):1045–72, 2003.

[31] George W. Evans and Seppo Honkapohja. Expectations and the stability problem for optimal monetary policy design. *Review of Economic Studies*, 70:807–824, 2003.

[32] George W Evans, Seppo Honkapohja, and Noah Williams. Generalized stochastic gradient learning. *CESifo Working Papers*, No. 1576, 2005.

[33] Claude Fourgeaud, Christian Gourieroux, and Jacqueline Pradel. Learning procedures and convergence to rationality. *Econometrica*, 54(4):845–68, 1986.

[34] M I Freidlin and A D Wentzell. *Random Perturbations of Dynamical Systems*. Springer-Verlag, New York, second edition, 1998.

[35] G. Fuchs. Is error learning behavior stabilizing? *Journal of Economic Theory*, 20:300–317, 1979.

[36] G. Fuchs and G. Laroque. Dynamics of temporary equilibria and expectations. *Econometrica*, 44:1157–1178, 1976.

[37] Chryssi Giannitsarou. Heterogeneous learning. *Review of Economic Dynamics*, 6:885–906, 2003.

[38] Chryssi Giannitsarou. E–stability does not imply learnability. *Macroeconomic Dynamics*, 9:276–287, 2005.

[39] J.-M. Grandmont. On endogenous competitive business cycles. *Econometrica*, 53:995–1045, 1985.

[40] J.-M. Grandmont. Expectations formation and stability of large socioeconomic systems. *Econometrica*, 66:741–781, 1998.

[41] J.-M. Grandmont and G. Laroque. Stability of cycles and expectations. *Journal of Economic Theory*, 40:138–151, 1986.

[42] Seppo Honkapohja and Kaushik Mitra. Performance of monetary policy rules with internal central bank forecasting. *Journal of Economic Dynamics and Control*, 29:627–658, 2005.

[43] Seppo Honkapohja and Kaushik Mitra. Learning stability in economies with heterogeneous agents. *Review of Economic Dynamics*, 9(2):284–309, 2006.

[44] Charles R. Johnson. Sufficient conditions for d-stability. *Journal of Economic Theory*, 9:53–62, 1974.

[45] Michihiro Kandori, George J Mailath, and Rafael Rob. Learning, mutation, and long run equilibria in games. *Econometrica*, 61:29–56, 1993.

[46] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, Berlin and Heidelberg, 1991.

[47] Kenneth Kasa. Learning, large deviations, and recurrent currency crises. *International Economic Review*, 45:141–173, 2004.

[48] R. Z. Khasminskii. *Stochastic Stability of Differential Equations*. Sijthoff & Noordhoof, Alphen aan den Rijn, The Netherlands, 1980.

[49] Alan. P. Kirman and Nicolaas J. Vriend. Evolving market structure: An ACE model of price dispersion and loyalty. *Journal of Economic Dynamics and Control*, 25:459–502, 2001.

[50] Dmiti Kolyuzhnov, Anna Bogomolova, and Sergey Slobodyan. Escape dynamics: A continuous–time approximation. CERGE–EI Working Paper 285, 2006.

[51] Dmitri Kolyuzhnov. Economic dynamics under heterogeneous learning: Necessary and sufficient conditions for stability. Manuscript, CERGE-EI, 2006.

[52] H. J. Kushner. Robustness and approximation of escape times and large deviations estimates for systems with small noise effects. *SIAM Journal of Applied Mathematics*, 44:160–182, 1984.

[53] Finn E. Kydland and Edward C. Prescott. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy*, 85:473–491, 1977.

[54] Helmut Lütkepohl. *Handbook of Matrices*. John Wiley & Sons, Chichester, 1996.

[55] Albert Marcet and Thomas J Sargent. Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory*, 48(2):337–68, 1989.

[56] Ramon Marimon. Learning from learning in economics. In David M Kreps and Kenneth F Wallis, editors, *Advances in Economics and Econometrics: Theory and Applications.*, volume 1, chapter 9, pages 278–315. Cambridge University Press, Cambridge, 1997.

[57] B. T. McCallum. On nonuniqueness in linear rational expectations models: An attempt at perspective. *Journal of Monetary Economics*, 11:134–168, 1983.

[58] Athanasios Orphanides. Inflation targeting under imperfect knowledge. FRB San Francisco Working Paper 2006-14, 2006.

[59] R V Roy. Large deviations theory, weak–noise asymptotics nd first–passage problems: Review and results. In M Leamaire, J.-L. Favre, and A Mebarki, editors, *Applications of Statistics and Probability*, pages 1129–1135. A. A. Balkema, Rotterdam, 1995.

[60] Thomas J Sargent. *Bounded Rationality in Macroeconomics*. Oxford University Press, Clarendon Press, Oxford and New York, 1993.

[61] Thomas J Sargent. *The Conquest of American Inflation*. Princeton University Press, 1999.

[62] Lars E. O. Svensson. Inflation targeting as a monetary policy rule. *Journal of Monetary Policy*, 43:607–654, 1999.

[63] C. E. Walsh. *Monetary Theory and Policy*. MIT Press, 2 edition, 2003.

[64] Noah Williams. *Escape Dynamics in Learning Models*. PhD thesis, University of Chicago, 2001.

[65] Noah Williams. Stability and long run equilibrium in stochastic fictitious play. Manuscript, Princeton University, 2002.

[66] Noah Williams. Adaptive learning and business cycles. Manuscript, Princeton University, 2003.

[67] Noah Williams. Small noise asymptotics for a stochastic growth model. *Journal of Economic Theory*, 119(2):271–298, 2004.

[68] Michael Woodford. Control of the public debt: A requirement for price stability. *NBER Working Paper Series*, w.p. 5684, 1996.

[69] Michael Woodford. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, 2003.

[70] J Worms. Moderate deviations for stable markov chains and regression models. *Electronic Journal of Probability*, 4(8):1–28, 1999.

Appendix A

Appendix to Chapter 2

A.1 Large deviations theory

Definition A.1 ***Action functional*** *for diffusion* $d\varphi_t = A\varphi_t dt + \sqrt{\epsilon} B dW_t$ *is defined as*

$$I_{0T}(\varphi) = \inf_{\left\{\dot{\varphi}_t = A\varphi_t + B\dot{g}_t\right\}} \tfrac{1}{2}\int_0^T \left|\dot{g}_t\right|^2 dt \text{ (see Dembo and Zeitouni [22, p. 214]).}$$

The results on the mean exit time and dominant escape point are given in Dembo and Zeitouni [22, Theorem 5.7.11].

Consider the system $dx_t^\varepsilon = b\left(x_t^\varepsilon\right) dt + \sqrt{\epsilon}\sigma\left(x_t^\varepsilon\right) dW_t,\ \ x_t^\varepsilon \in \mathbb{R}^d,\ x_0^\varepsilon = x.$

Assumption A.1 *The unique stable equilibrium point in* D *(open, bounded domain) of the d-dimensional ODE* $\dot{x}_t = b\left(x_t\right)$ *is at* $O \in D$ *and* $x_0 \in D \Longrightarrow \forall t > 0, x_t \in D$ *and* $\lim_{t\longrightarrow\infty} x_t = O$.

Assumption A.2 *All the trajectories of the deterministic ODE* $\dot{x}_t = b\left(x_t\right)$ *starting at* $x_0 \in \partial D$ *converge to* O *as* $t \longrightarrow \infty$.

Assumption A3 $\overline{I} \overset{def}{=} \inf_{y\in\partial D} I\left(O, y\right) < \infty.$

Assumption A.4 *There exists an* $M < \infty$ *such that for all* $\rho > 0$ *small enough and all* x, y *with* $|x - z| + |y - z| \leq \rho$ *for some* $z \in \partial D \cup \{O\}$, *there is a function* u *satisfying that* $||u|| < M$ *and* $\varphi_{T(\rho)} = y$, *where* $\varphi_t = x + \int_0^t b\left(\varphi_s\right) ds + \int_0^t \sigma\left(\varphi_s\right) u_s ds$ *and* $T\left(\rho\right) \longrightarrow 0$ *as* $\rho \longrightarrow 0$.

Definition A.2 $\tau^\varepsilon \overset{def}{=} \inf\left\{t > 0 \colon x_t^\varepsilon \in \partial D\right\}.$

Theorem A.1 *(Dembo and Zeitouni [22, Theorem 5.7.11]) Assume A.1–A.4. (a) For all* $x \in D$ *and all* $\delta > 0$, $\lim_{\varepsilon\longrightarrow 0} P_x\left(e^{\left(\overline{I}+\delta\right)/\varepsilon} > \tau^\varepsilon > e^{\left(\overline{I}-\delta\right)/\varepsilon}\right) = 1$. *Moreover, for all* x, $\lim_{\varepsilon\longrightarrow 0} \varepsilon \ln E_x\left(\tau^\varepsilon\right) = \overline{I}$. *(b) If* $N \subset \partial D$ *is a closed set and* $\inf_{z\in N} I\left(O, z\right) > \overline{I}$, *then for any* $x \in D$, $\lim_{\varepsilon\longrightarrow 0} P_x\left(x_{\tau_\varepsilon}^\varepsilon \in N\right) = 0$. *In particular, if there exists* $z^* \in \partial D$ *such that* $I\left(O, z^*\right) < I\left(O, z\right)$ *for all* $z \neq z^*, z \in \partial D$ *, then* $\forall\delta > 0, \forall x \in D$, $\lim_{\varepsilon\longrightarrow 0} P_x\left(\left|x_{\tau_\varepsilon}^\varepsilon - z^*\right|\right) = 1$.

Part a) of the theorem characterizes the escape probability and the mean escape time, and part b) gives the dominant escape point.

A.2 Minimizing the action functional

We need to solve the problem

$$\min \frac{1}{2} \int_0^T \|u_t\|^2 \, dt,$$

subject to

$$\begin{aligned} \dot{\varphi}_t &= A\varphi_t + Bu_t, \\ \varphi(0) &= 0, \; \varphi(T) \in \partial D \end{aligned}$$

We know that matrix B is singular, and therefore, the system (A, B) is not reachable. Such systems are usually converted into the *standard form* as follows.[1] Change the coordinates so that $\varphi = Tz$:

$$z = T^{-1}\varphi = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix},$$

where dimension of z_1 is r, a dimension of the reachable subspace. In the new coordinates, we get

$$\begin{aligned} T\dot{z} &= ATz + Bu, \text{ or} \\ \dot{z} &= \underbrace{T^{-1}AT}_{\bar{A}}z + \underbrace{T^{-1}B}_{\bar{B}}u. \end{aligned}$$

The matrix T is constructed as $[T_1 \,|T_2\,]$, where T_1 consists of columns that form the basis of the reachable subspace. (It is convenient to select columns of T to be the [orthonormal] basis of Gramian G in the initial problem. T_1 are the columns corresponding to the nonzero eigenvalues of G; therefore, they constitute the basis of reachable subspace.) Columns of T_2 form the (orthonormal) basis of the complement to the reachable subspace. By construction, matrix T is invertible.

Let us find the structure of $\bar{A}$ and $\bar{B}$. Look at $AT = A\,[T_1\,|T_2\,] = [T_1\,|T_2\,]\,\overline{A}$. Reachable space is invariant for all controls including $u = 0$; therefore, the range of $[T_1\,|T_2\,]\,\overline{A}$ should not include vectors from T_2. This could be achieved if

$$\overline{A} = \begin{bmatrix} \overline{A}_1 & \overline{A}_{12} \\ 0 & \overline{A}_2 \end{bmatrix}.$$

Similarly, no control should push the system out of reachable subspace; this means that

$$B = [T_1\,|T_2\,]\,\overline{B} = [T_1\,|T_2\,] \begin{bmatrix} \overline{B}_1 \\ 0 \end{bmatrix}.$$

[1] See details in Dahleh, Dahleh, and Verghese [21] and Boyd [8].

With this, we can now write our system as

$$\dot{z} = T^{-1}ATz + T^{-1}Bu = \begin{bmatrix} \overline{A}_1 & \overline{A}_{12} \\ 0 & \overline{A}_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} \overline{B}_1 \\ 0 \end{bmatrix} u.$$

As we are interested in the movement from initial point $z(0) = 0$, unreachable dimensions z_2 cannot influence dynamics of z_1 :

$$\dot{z}_1 = \overline{A}_1 z_1 + \overline{B}_1 u.$$

This system is called a *reachable subsystem* of the original one.

Let us find $\bar{A}_1$ and $\bar{B}_1$. $T^{-1}AT = \bar{A}$, or $\begin{bmatrix} T_1' \\ T_2' \end{bmatrix} [A] [T_1 | T_2] = \begin{bmatrix} \overline{A}_1 & \overline{A}_{12} \\ 0 & \overline{A}_2 \end{bmatrix}$.

Therefore, $T_1' A T_1 = \overline{A}_1$. For $\overline{B}_1$, $T^{-1}B = \bar{B} \Rightarrow \begin{bmatrix} T_1' \\ T_2' \end{bmatrix} B = \begin{bmatrix} \overline{B}_1 \\ 0 \end{bmatrix} \Longrightarrow T_1' B = \overline{B}_1$.

In the new variables our problems transform into

$$\begin{aligned} \min I_{0T} &= \frac{1}{2}\int_0^T \|u_t\|^2\, dt, \\ & \text{subject to} \\ \dot{z}_1 &= \overline{A}_1 z_1 + \overline{B}_1 u, \\ z_1(0) &= 0,\ T_1 z_1(T) \in \partial D. \end{aligned}$$

(Note that z_2 stays zero under our dynamics; therefore, $Tz = [T_1 | T_2] \begin{bmatrix} z_1 \\ 0 \end{bmatrix} = T_1 z_1$).

This problem is easily solved as the system $(\overline{A}_1, \overline{B}_1)$ is reachable by construction. The standard result is that $I = \frac{1}{2} z_{1,des}^T \cdot \overline{G}^{-1} \cdot z_{1,des}$, where $\overline{G}$ is Gramian in the reduced problem, given as a solution of the matrix Lyapunov equation $\overline{A}_1\overline{G} + \overline{GA}_1' + \overline{B}_1\overline{B}_1' = 0$ and $T_1 z_{1,des} \in \partial D$.

In the case when the set ∂D is the surface of the cylinder — a sphere of radius R in γ space,$\|\gamma\| = R$, and no binding restrictions in space of components of R — the problem of minimizing the action functional becomes

$$\begin{aligned} &\min \frac{1}{2} z_1^T \cdot \overline{G}^{-1} \cdot z_1, \\ &\text{s. t.} \quad \left(I_{27}^6 \overline{T}_1 z_1\right)^T \cdot \left(I_{27}^6 \overline{T}_1 z_1\right) = R^2, \end{aligned}$$

where I_{27}^6 is $27 \times 27 zero$ matrix 6×6 identity matrix in the upper left corner. After defining $v = \overline{G}^{-1/2} \cdot z_1$, it is straightforward to get solution $\varphi_{des} = \pm \frac{R}{\lambda_1} T_1 \overline{G}^{1/2} \xi$, where ξ is the unit eigenvector of matrix $\Gamma = \overline{G}^{1/2}\overline{T}_1'\overline{T}_1\overline{G}^{1/2}$, corresponding to the largest eigenvalue, $(\lambda_1)^2$. Note that if matrix Σ is block–diagonal, as in the static model, the eigenvector ξ coincides with the "largest" eigenvector of $\overline{G}$, and eigenvalue λ_1 coincides with the largest eigenvalue of $\overline{G}$.

For the problem when ∂D is given numerically, one has to find $\min\limits_{z_1 \in \partial D} \frac{1}{2} z_1^T \cdot \overline{G}^{-1} \cdot z_1$. All the points on the boundary are given parametrically by 2–dimensional function $\varepsilon(t)$, where t is the index number of the point in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space. Write $Fz_1 = \varepsilon(t)$, where F transforms 13 dimensions of z_1 into 2–dimensional space $(\tilde{\gamma}_1, \tilde{\gamma}_2)$. Define $v = \overline{G}^{-\frac{1}{2}} z_1$. Then the problem becomes

$$\begin{aligned} & \min \frac{1}{2} \|v\|^2 \\ & \text{s. t.} \\ F\overline{G}^{\frac{1}{2}} v \; = \; & \varepsilon(t) \end{aligned}$$

and the solution of this problem is

$$v_t = pinv\left(F\overline{G}^{\frac{1}{2}}\right)\varepsilon(t), \bar{t} = \arg\min \left\| pinv\left(F\overline{G}^{\frac{1}{2}}\right)\varepsilon(t)\right\|;$$

$\varepsilon(\bar{t})$ is the predicted point of escape in $(\tilde{\gamma}_1, \tilde{\gamma}_2)$ space. To transform this point into the original 27–dimensional space of beliefs, use the following transformation:

$$z_{1\bar{t}} = \overline{G}^{\frac{1}{2}} pinv\left(F\overline{G}^{\frac{1}{2}}\right)\varepsilon(\bar{t}) \Longrightarrow \varphi = T_1'\overline{G}^{\frac{1}{2}} pinv\left(F\overline{G}^{\frac{1}{2}}\right)\varepsilon(\bar{t}).$$

For the problem (disregarding the mean dynamics), we set $A = 0$ and using the general result above, get $I = \frac{1}{2} z_{1,des}^T \cdot \overline{G}^{-1} \cdot z_{1,des}$. In this case $\overline{G}$ is defined for arbitrary time T as $\overline{G}^{-1} = \left(\overline{B}_1\overline{B}_1'\right)^{-1}\frac{1}{T} = (T_1'BB'T_1)^{-1}\frac{1}{T} = T_1(BB')^{-1}T_1'\frac{1}{T} = T_1\left(\Sigma\left(\bar{\theta}\right)\right)^{-1}T_1'\frac{1}{T}$,where T_1 is the basis of spectral decomposition of $\Sigma\left(\bar{\theta}\right)$, and in the same time the orthonormal basis of the reachable subspace. For any T the solution of the problem on the cylinder is expressed by the formula for escape out of the cylinder written above, where instead of $\overline{G}$ one uses $T_1\Sigma\left(\bar{\theta}\right)T_1'$. In the model of CWS the resulting direction almost coincides with the eigenvector corresponding to the largest eigenvalue of $\Sigma\left(\bar{\theta}\right)$.

A.3 Formula for the third way of deriving mean escape time

The third way of deriving escape dynamics characteristics is based on the "modified" continuous–time approximation without drift term $d\varphi_t = \sqrt{\epsilon}\Sigma^{1/2}(\bar{\theta})dW_t$. To find the projection of the process on the most probable direction of escape $\tilde{v}_1$, the "largest" eigenvector of Σ, we multiply the above expression by this eigenvector from the left. The resulting diffusion is $d\varphi_t^{projection} = \sqrt{\epsilon\lambda} dW_t$, where λ is the largest eigenvalue of Σ. Then we use the formula for the mean exit time for one–dimensional Brownian motion in Karatzas and Shreve [46, Eq. 5.62, p. 345]. For a process $Y_t = x + \int_0^t \sigma(Y_s)\, dW_s$, the mean of exit time $T_{a,b}(x) = \inf\{t \geq 0; Y_t \notin (a,b)\}$ is expressed as $ET_{a,b}(x) = \int_a^b \frac{(\min(x,y)-a)(b-\max(x,y))}{b-a} \cdot \frac{2dy}{\sigma^2(y)}$. In our case x, the starting point of the projection of the process of deviations from the SCE, is zero, $\sigma(y)$ is replaced by $\sqrt{\epsilon\lambda}$, and the interval (a,b) is given by $(-rad, rad)$, where rad is the distance between the SCE and the point where the "largest" eigenvector of Σ crosses the cylinder used in the first and third way of deriving escape dynamics. After plugging these values into the expression for the mean exit time and evaluating the integral, we get the formula for the mean escape time: $E\tau^{\varepsilon} = \frac{rad^2}{\epsilon\lambda}$.

Appendix B

Appendix to Chapter 4

Here I provide the reader with definitions and theorems adapted from mathematics literature that I used for deriving conditions for δ–stability. These results are structured according to the approach which is used for deriving stability conditions.

B.1 General definition of stability and D–stability of a matrix

Definition B.1 *Matrix A is stable if all the solutions of the system of ordinary differential equations $\dot{x}(t) = Ax(t)$ converge toward zero as t converges to infinity.*

Theorem B.1 *Matrix A is stable if and only if all its eigenvalues have negative real parts.*

Definition B.2 *($D-$stability) Matrix A is D–stable if DA is stable for any positive diagonal matrix D.*

B.2 Lyapunov theorem approach

Theorem B.2 *(Lyapunov) A real $n \times n$ matrix A is a stable matrix if and only if there exists a positive definite matrix H such that $A'H + HA$ is negative definite.*

Theorem B.3 *(Arrow-McManus, 1958) Matrix A is D–stable if there exists a positive diagonal matrix C such that $A'C + CA$ is negative definite.*

B.3 Negative diagonal dominance approach

Definition B.3 *(introduced by McKenzie) A real $n \times n$ matrix A is dominant diagonal if there exist n real numbers $d_j > 0, j = 1, ..., n$, such that $d_j|a_{jj}| > \sum d_i|a_{ij}| : i \neq j)$, $j = 1, \ldots, n$ This is called "column" diagonal dominance. "Row" diagonal dominance is defined as the existence of $d_i > 0$ such that $d_i|a_{ii}| > \sum d_j|a_{ij}| : j \neq i), i = 1, \ldots, n$.*

Theorem B.4 *(sufficient condition for stability, McKenzie, 1960): If an $n \times n$ matrix A is dominant diagonal and its diagonal is composed of negative elements ($a_{ii} < 0$, all $i = 1, \ldots, n$), then the real parts of all its eigenvalues are negative, i.e., A is stable.*

Corollary B.1 *If A has negative diagonal dominance, then it is D–stable.*

B.4 Characteristic equation approach

Theorem B.5 *(Routh-Hurwitz necessary and sufficient conditions for negativity of eigenvalues of a matrix) Consider the following characteristic equation*

$$|\lambda I - A| = \lambda^n + b_1\lambda^{n-1} + ... + b_{n-1}\lambda + b_n = 0$$

determining n eigenvalues λ of a real $n \times n$ matrix A, where I is the identity matrix. Then eigenvalues λ all have negative real parts if and only if $\Delta_1 > 0, \Delta_2 > 0, ..., \Delta_n > 0$, where

$$\Delta_k = \begin{vmatrix} b_1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ b_3 & b_2 & b_1 & 1 & 0 & \cdots & 0 \\ b_5 & b_4 & b_3 & b_2 & b_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ b_{2k-1} & b_{2k-2} & b_{2k-3} & b_{2k-4} & b_{2k-5} & \cdots & b_k \end{vmatrix}.$$

B.5 Alternative definition of $D-$stability approach

Theorem B.6 *(From Observation (iv) in Johnson [44]). Consider $M_n(C)$, the set of all complex $n \times n$ matrices, and D_n, the set of all $n \times n$ diagonal matrices with positive diagonal entries. Take $A \in M_n(C)$ and suppose that there is an $F \in D_n$ such that FA is stable. Then A is D–stable if and only if $A \pm iD$ is non–singular for all $D \in D_n$. If $A \in M_n(R)$, the set of all $n \times n$ real matrices, then "$\pm$" in the above condition may be replaced with "$+$" since, for a real matrix, any complex eigenvalues come in conjugate pairs.*

B.6 Proofs of propositions in Chapter 4

B.6.1 Proof of Proposition 4.2 (The criterion for stability of a structurally heterogeneous economy under mixed RLS/SG learning for the diagonal environment case under any (possibly different) degrees of inertia of agents, $\delta > 0$)

We have to consider conditions for stability for any positive $(\delta_1, ..., \delta_S)$ of the following matrices

$$D_1\Omega = \begin{pmatrix} \delta_1 I_n & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \delta_S I_n \end{pmatrix} \begin{pmatrix} A_1 - I_n & \cdots & A_S \\ \vdots & \ddots & \vdots \\ A_1 & \cdots & A_S - I_n \end{pmatrix}$$

and

$$D_w\Omega_F = \begin{pmatrix} D_{w1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{wS} \end{pmatrix} \begin{pmatrix} F' \otimes A_1 - I_{nk} & \cdots & F' \otimes A_S \\ \vdots & \ddots & \vdots \\ F' \otimes A_1 & \cdots & F' \otimes A_S - I_{nk} \end{pmatrix},$$

where $\begin{matrix} D_{wh} = \delta_h I_{nk}, h = \overline{1, S_0} \\ D_{wh} = \delta_h \left(M_w \otimes I_n\right), h = \overline{S_0 + 1, S} \end{matrix}$, $F = diag(\rho_1, ..., \rho_k), M_w = diag\left(\frac{\sigma_1^2}{1-\rho_1^2}, ..., \frac{\sigma_k^2}{1-\rho_k^2}\right)$.

The expression for $D_w\Omega_F$ in the diagonal case looks as follows

$$D_w\Omega_F = \begin{pmatrix} D_{w1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_{wS} \end{pmatrix} \begin{pmatrix} F' \otimes A_1 - I_{nk} & \cdots & F' \otimes A_S \\ \vdots & \ddots & \vdots \\ F' \otimes A_1 & \cdots & F' \otimes A_S - I_{nk} \end{pmatrix} =$$

$$= \; diag(\underbrace{\delta_1, \ldots, \delta_1}_{nk}, \ldots, \underbrace{\delta_{S_0}, \ldots, \delta_{S_0}}_{nk}, \underbrace{\tfrac{\delta_{S_0+1}\sigma_1^2}{1-\rho_1^2}, \ldots, \tfrac{\delta_{S_0+1}\sigma_1^2}{1-\rho_1^2}}_{n}, \ldots, \underbrace{\tfrac{\delta_{S_0+1}\sigma_k^2}{1-\rho_k^2}, \ldots, \tfrac{\delta_{S_0+1}\sigma_k^2}{1-\rho_k^2}}_{n}, \ldots$$

$$\ldots, \underbrace{\tfrac{\delta_S\sigma_1^2}{1-\rho_1^2}, \ldots, \tfrac{\delta_S\sigma_1^2}{1-\rho_1^2}}_{n}, \ldots, \underbrace{\tfrac{\delta_S\sigma_k^2}{1-\rho_k^2}, \ldots, \tfrac{\delta_S\sigma_k^2}{1-\rho_k^2}}_{n}) \times$$

$$\times \begin{pmatrix} \rho_1 A_1 - I_n & \cdots & 0 & \cdots & \rho_1 A_S & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_k A_1 - I_n & \cdots & 0 & \cdots & \rho_k A_S \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \rho_1 A_1 & \cdots & 0 & \cdots & \rho_1 A_S - I_n & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_k A_1 & \cdots & 0 & \cdots & \rho_k A_S - I_n \end{pmatrix}.$$

After some permutations of rows and columns that do not change the absolute value of the determinant of $D_w\Omega_F - \mu I$, I obtain that the following characteristic equation for eigenvalues μ of $D_w\Omega_F$

$$\det\left[D_w\Omega_F - \mu I\right] = 0$$

is equivalent to

$$\begin{aligned} 0 \;=\; & \det[diag((\underbrace{\delta_1,\ldots,\delta_1}_{n},\ldots,\underbrace{\delta_{S_0},\ldots,\delta_{S_0}}_{n},\underbrace{\tfrac{\delta_{S_0+1}\sigma_1^2}{1-\rho_1^2},\ldots,\tfrac{\delta_{S_0+1}\sigma_1^2}{1-\rho_1^2}}_{n},\ldots,\underbrace{\tfrac{\delta_S\sigma_1^2}{1-\rho_1^2},\ldots,\tfrac{\delta_S\sigma_1^2}{1-\rho_1^2}}_{n}),\ldots \\ & \ldots,(\underbrace{\delta_1,\ldots,\delta_1}_{n},\ldots,\underbrace{\delta_{S_0},\ldots,\delta_{S_0}}_{n},\underbrace{\tfrac{\delta_{S_0+1}\sigma_k^2}{1-\rho_k^2},\ldots,\tfrac{\delta_{S_0+1}\sigma_k^2}{1-\rho_k^2}}_{n},\ldots,\underbrace{\tfrac{\delta_S\sigma_k^2}{1-\rho_k^2},...,\tfrac{\delta_S\sigma_k^2}{1-\rho_k^2}}_{n}))\times \end{aligned}$$

$$\times diag(\begin{bmatrix} \rho_1 A_1 - I_n - \frac{\mu I_n}{\delta_1} & \cdots & \rho_1 A_S \\ \vdots & \ddots & \vdots \\ \rho_1 A_1 & \cdots & \rho_1 A_S - I_n - \frac{(1-\rho_1^2)\mu I_n}{\delta_S\sigma_1^2} \end{bmatrix},\ldots$$

$$\ldots,\begin{bmatrix} \rho_k A_1 - I_n - \frac{\mu I_n}{\delta_1} & \cdots & \rho_k A_S \\ \vdots & \ddots & \vdots \\ \rho_k A_1 & \cdots & \rho_k A_S - I_n - \frac{(1-\rho_k^2)\mu I_n}{\delta_S\sigma_k^2} \end{bmatrix})],$$

or, in matrix form:

$$0 = \det\begin{bmatrix} \tilde{D}_1\Omega_{\rho_1} - \mu I_{nS} & \\ & \\ & \tilde{D}_k\Omega_{\rho k} - \mu I_{nS} \end{bmatrix} = \prod_{l=1}^{k} \det\left[\tilde{D}_l\Omega_{\rho_l} - \mu I_{nS}\right],$$

where

$$\begin{aligned} \tilde{D}_l &= \begin{pmatrix} \delta_1 I_n & & & \cdots & 0 \\ & \ddots & & & \\ & & \delta_{S_0} I_n & & \\ \vdots & & & \frac{\delta_{S_0+1}\sigma_l^2}{1-\rho_l^2} I_n & \vdots \\ & & & \ddots & \\ 0 & & & \cdots & \frac{\delta_S\sigma_l^2}{1-\rho_l^2} I_n \end{pmatrix}, \Omega_{\rho_l} = \begin{pmatrix} \rho_l A_1 - I_n & \cdots & \rho_l A_S \\ \vdots & \ddots & \vdots \\ \rho_l A_1 & \cdots & \rho_l A_S - I_n \end{pmatrix}, \\ l &= \overline{1,k}. \end{aligned}$$

Thus, the analysis of stability of $D_w\Omega_F$, defined in (4.11), is equivalent to the analysis of stability of $\tilde{D}_l\Omega_{\rho_l}$, $\forall l = \overline{1,k}$.

So, the analysis of the stability of $D_w\Omega_F$ can be split into the analysis of the stability of the unrelated matrix blocks. Changing notation $\delta_h := \frac{\delta_h\sigma_l^2}{1-\rho_l^2} > 0$ for $h = \overline{S_0+1, S}$ for each case $l = \overline{1,k}$, I obtain that the analysis of stability of $D_w\Omega_F$ for any $\delta > 0$ is equivalent to the analysis of stability of k matrices $D_1\Omega_{\rho_l}$. Introducing notation $\rho_0 = 1$, I can write the general criterion for stability of a structurally heterogeneous economy under mixed RLS/SG learning for the diagonal environment case under any (possibly different) degrees of inertia of agents, $\delta > 0$ as follows: $D_1\Omega_{\rho_l}$ is stable for all $l = 0, 1, ..., k$. *Q.E.D.*

B.6.2 Proof of Proposition 4.3

Use "columns" negative diagonal dominance of Ω_{ρ_l}, which is sufficient for the real parts of eigenvalues of $D_1\Omega_{\rho_l}$ to be negative; look for a condition which would be sufficient for negative diagonal dominance in this setup. As weights for rows use $(\phi_1(\psi_1, ..., \psi_n), ..., \phi_s(\psi_1, ..., \psi_n))$, $\phi_i > 0$, $\psi_h > 0$, $\sum\limits_i \psi_i = 1, \sum\limits_h \phi_h = 1$.

For any l take any block h and any column j

$$\begin{cases} \rho_l a_{jj}^h - 1 < 0 \text{ - negative diagonal} \\ \phi_h\psi_j \left|\rho_l a_{jj}^h - 1\right| > (\phi_1 + ... + \phi_s) \sum\limits_i \psi_i \left|\rho_l a_{ij}^h\right| - \phi_h\psi_j \left|\rho_l a_{jj}^h\right| \text{ - dominance} \end{cases} \forall j, \forall h, \forall l$$

$$\Updownarrow$$

$$\begin{cases} \rho_l a_{jj}^h - 1 < 0 \\ -\phi_h\psi_j\rho_l a_{jj}^h + \phi_h\psi_j > (\phi_1 + ... + \phi_s) \sum\limits_i \psi_i \left|\rho_l a_{ij}^h\right| - \phi_h\psi_j \left|\rho_l a_{jj}^h\right| \end{cases} \forall j, \forall h, \forall l$$

$$\Updownarrow$$

$$\text{Case 1}\begin{cases} 0 \le \rho_l a_{jj}^h < 1 \\ \sum\limits_i \psi_i \left|\rho_l a_{ij}^h\right| < \frac{\phi_h\psi_j}{\underbrace{\phi_1 + ... + \phi_S}_{=1}} \end{cases} \forall j, \forall h, \forall l$$

$$\cup$$

$$\text{Case 2}\begin{cases} \rho_l a_{jj}^h < 0 \\ \sum\limits_i \psi_i \left|\rho_l a_{ij}^h\right| < \frac{\phi_h\psi_j}{\underbrace{\phi_1 + ... + \phi_S}_{=1}} - \frac{2\varphi_h\psi_j}{\underbrace{\phi_1 + ... + \phi_S}_{=1}}\rho_l a_{jj}^h \end{cases} \forall j, \forall h, \forall l$$

Since in the second case $\rho_l a_{jj}^h < 0$, one may formulate the following sufficient condition $\sum\limits_i \psi_i \left|\rho_l a_{ij}^h\right| < \phi_h\psi_j$ $\forall j, \forall h, \forall l$. The condition $1 > \rho_l a_{jj}^h$ is implied by this relation, and the condition of case 2 is also satisfied. To prove that $1 > \rho_l a_{jj}^h$, notice that

$$\sum\limits_i \psi_i \left|\rho_l a_{ij}^h\right| < \phi_h\psi_j \Longrightarrow \underbrace{\frac{\sum\limits_{i\neq j} \psi_i \left|\rho_l a_{ij}^h\right|}{\psi_j}}_{>0} + \underbrace{\left|\rho_l a_{jj}^h\right|}_{>0} < \phi_h < 1 \Longrightarrow \left|\rho_l a_{jj}^h\right| < 1 \Longrightarrow \rho_l a_{jj}^h < 1.$$

As $|\rho_l| < 1$, the derived sufficient condition follows from $\sum\limits_i \psi_i \left|a_{ij}^h\right| < \phi_h\psi_j$ $\forall j, \forall h,$

that is, the condition for $l = 0$ ($\rho_0 = 1$). So this condition alone is sufficient for $\delta-$stability. This is the condition of Proposition 4.2. $Q.E.D.$

B.6.3 Proof of Proposition 4.4

1. $\beta^{AG\,\mathrm{mod}}(\psi,\phi)\big|_{\substack{\phi-any\\ \psi-any}} = S\sum_h \phi_h \sum_i \psi_i \sum_j \left|a_{ij}^h\right| \leq$

$$\leq S\sum_h \phi_h \sum_j \sum_i \psi_i \max_{h,i}\left|a_{ij}^h\right| = S\sum_j \underbrace{\left(\sum_h \sum_i \phi_h \psi_i\right)}_{=1} \max_{h,i}\left|a_{ij}^h\right| = \beta_1^{AG\max}.$$

2. $\beta^{AG\,\mathrm{mod}}(\psi,\phi)\big|_{\substack{\phi-any\\ \psi=\frac{1}{S}}} = S\sum_h \underbrace{\frac{1}{S}}_{\phi_h} \sum_i \psi_i \sum_j \left|a_{ij}^h\right| = \sum_h \sum_i \psi_i \sum_j \left|a_{ij}^h\right| \leq$

$$\leq \underbrace{\left(\sum_i \psi_i\right)}_{=1} \max_i \sum_h \sum_j \left|a_{ij}^h\right| = \beta_2^{AG\max}$$

3. $\beta^{AG\,\mathrm{mod}}(\psi,\phi)\big|_{\substack{\phi=\frac{1}{n}\\ \psi-any}} = S\sum_h \phi_h \sum_i \underbrace{\frac{1}{n}}_{\psi_i} \sum_j \left|a_{ij}^h\right| \leq S\sum_i \frac{1}{n}\sum_h \sum_j \phi_h \max_{h,j}\left|a_{ij}^h\right| =$

$$= S\sum_i \frac{1}{n}\max_{h,j}\left|a_{ij}^h\right| \underbrace{\left(\sum_h \sum_j \phi_h\right)}_{=n} = S\sum_i \max_{h,j}\left|a_{ij}^h\right| = \beta_3^{AG\max}$$

4. $\beta^{AG\,\mathrm{mod}}(\psi,\phi)\big|_{\substack{\phi=\frac{1}{n}\\ \psi=\frac{1}{S}}} = S\sum_h \underbrace{\frac{1}{S}}_{\phi_h} \sum_i \underbrace{\frac{1}{n}}_{\psi_i} \sum_j \left|a_{ij}^h\right| = \sum_h \sum_i \frac{1}{n}\sum_j \left|a_{ij}^h\right| \leq$

$$\leq \sum_h \frac{1}{n}\sum_j \max_j \sum_i \left|a_{ij}^h\right| = \sum_h \max_j \sum_i \left|a_{ij}^h\right| \underbrace{\frac{1}{n}\sum_j 1}_{=1} = \beta_4^{AG\max}, Q.E.D.$$

B.6.4 Proof of Proposition 4.5

1. for $\beta_1^{AG\max}$:

We have $\beta_1^{AG\max} = S\sum_j \max_{h,i}\left|a_{ij}^h\right| < 1$ and have to prove that there exist weights ψ and ϕ such that $\frac{\sum_i \psi_i \left|a_{ij}^h\right|}{\psi_j} < \phi_h \ \forall j, \forall h$.

Let us take $\phi_h = \frac{1}{S}\ \forall h$, and $\psi_j = S\max_{h,i}\left|a_{ij}^h\right| + \frac{\overbrace{1 - S\sum_j \max_{h,i}\left|a_{ij}^h\right|}^{>0}}{n}\ \forall j$. These can be considered as weights since $\sum_{h=1}^{S}\phi_h = 1, 0 < \phi_h < 1$ and $\sum_{j=1}^{n}\psi_j = 1, 0 < \psi_j < 1$.

Notice that $\frac{\psi_j}{S} > \max\limits_{h,i} \left|a_{ij}^h\right| = \sum\limits_i \psi_i \max\limits_{h,i} \left|a_{ij}^h\right| > \sum\limits_i \psi_i \left|a_{ij}^h\right|, \forall j, \forall h$, or, after rewriting: $\sum\limits_i \psi_i \left|a_{ij}^h\right| < \psi_j \underbrace{\phi_h}_{=\frac{1}{S}}, \forall j, \forall h.$

4. for $\beta_4^{AG\max}$:

We have $\beta_4^{AG\max} = S \sum\limits_h \max\limits_j \sum\limits_i \left|a_{ij}^h\right| < 1$ and have to prove that there exist weights ψ and ϕ such that $\frac{\sum\limits_i \psi_i \left|a_{ij}^h\right|}{\psi_j} < \phi_h \ \forall j, \forall h$.

Let us take $\psi_j = \frac{1}{n} \ \forall h$, $\phi_h = \max\limits_j \sum\limits_i \left|a_{ij}^h\right| + \frac{\overbrace{1 - \sum\limits_h \max\limits_j \sum\limits_i \left|a_{ij}^h\right|}^{>0}}{S} \ \forall j$. These are weights as $\sum\limits_{h=1}^{S} \phi_h = 1, 0 < \phi_h < 1$ and $\sum\limits_{j=1}^{S} \psi_j = 1, 0 < \psi_j < 1$.

Notice that $\phi_h > \max\limits_j \sum\limits_i \left|a_{ij}^h\right| > \sum\limits_i \left|a_{ij}^h\right|, \forall j, \forall h$, or, after rewriting: $\frac{\sum\limits_i \overbrace{\psi_i}^{\frac{1}{n}} \left|a_{ij}^h\right|}{\underbrace{\psi_j}_{\frac{1}{n}}} = \sum\limits_i \psi_i \left|a_{ij}^h\right| < \phi_h, \forall j, \forall h.$

To prove the proposition for $\beta_2^{AG\max}$ and $\beta_3^{AG\max}$, I first derive a sufficient condition for δ – stability that follows from the "rows" diagonal dominance condition, which is also sufficient for stability of matrices $D_1\Omega_{\rho l}$. Therefore my derivation of this condition resembles the steps in the proof of Proposition 4.2. As weights for columns use $(d_1, ..., d_n, ..., d_1, ..., d_n)$, $d_i > 0$, $\sum\limits_i d_i = 1$

For any l take any block h and any row i.

$$\begin{cases} \rho_l a_{ii}^h - 1 < 0 \text{ - negative diagonal} \\ d_i \left|\rho_l a_{ii}^h - 1\right| > \sum\limits_h \sum\limits_j d_j \left|\rho_l a_{ij}^h\right| - d_i \left|\rho_l a_{ii}^h\right| \text{ - dominance} \end{cases} \forall i, \forall h, \forall l$$

$$\Updownarrow$$

$$\begin{cases} \rho_l a_{ii}^h - 1 < 0 \\ -d_i \rho_l a_{ii}^h + d_i > \sum\limits_h \sum\limits_j d_j \left|\rho_l a_{ij}^h\right| - d_i \left|\rho_l a_{ii}^h\right| \end{cases} \forall i, \forall h, \forall l$$

$$\Updownarrow$$

$$\text{Case 1} \begin{cases} 0 \le \rho_l a_{ii}^h < 1 \\ \sum\limits_h \sum\limits_j d_j \left|\rho_l a_{ij}^h\right| < d_i \end{cases} \forall i, \forall h, \forall l$$

$$\cup$$

$$\text{Case 2} \begin{cases} \rho_l a_{ii}^h < 0 \\ \sum\limits_h \sum\limits_j d_j \left|\rho_l a_{ij}^h\right| < d_i - 2 d_i \rho_l a_{ii}^h \end{cases} \forall i, \forall h, \forall l$$

Since in the second case $\rho_l a^h_{ii} < 0$, one may formulate the following sufficient condition $\sum_h \sum_j d_j \left|\rho_l a^h_{ij}\right| < d_i \ \forall i, \forall h, \forall l$. The condition $1 > \rho_l a^h_{ii}$ is implied by this relation, and the condition of case 2 is also satisfied. To prove that $1 > \rho_l a^h_{ii}$, notice that $\sum_h \sum_j d_j \left|\rho_l a^h_{ij}\right| < d_i \implies \underbrace{\sum_h \sum_{j \neq i} d_j \left|\rho_l a^h_{ij}\right|}_{>0} + \underbrace{\sum_h d_i \left|\rho_l a^h_{ii}\right|}_{>0} < d_i < 1 \implies \left|\rho_l a^h_{ii}\right| < 1 \implies \rho_l a^h_{ii} < 1$.

As $|\rho_l| < 1$, the derived sufficient condition follows from $\sum_h \sum_j d_j \left|a^h_{ij}\right| < d_i \ \forall i, \forall h$, that is, the condition for $l = 0$ ($\rho_0 = 1$). So this condition alone is sufficient for δ–stability.

Next I use the derived sufficient condition to prove Proposition 4.2 for $\beta^2_{\max}$ and $\beta^3_{\max}$.

2. for $\beta_2^{AG\max}$:

We have $\beta_2^{AG\max} = \max_i \sum_h \sum_j \left|a^h_{ij}\right| < 1$ and have to prove that there exist weights $d = (d_1, ..., d_n, ..., d_1, ..., d_n)$, $d_i > 0$, $\sum_i d_i = 1$, such that $\sum_h \sum_j d_j \left|a^h_{ij}\right| < d_i \ \forall i, \forall h$.

Let us take $d_j = \frac{1}{n} \ \forall j$.

Notice that $\sum_h \sum_j \left|a^h_{ij}\right| < \max_i \sum_h \sum_j \left|a^h_{ij}\right| < 1, \forall i, \forall h$, or, after rewriting: $\sum_h \sum_j \underbrace{\frac{1}{n}}_{d_j} \left|a^h_{ij}\right| < \underbrace{\frac{1}{n}}_{d_i}, \forall i, \forall h$.

3. for $\beta_3^{AG\max}$:

We have $\beta_3^{AG\max} = S \sum_h \max_{h,j} \left|a^h_{ij}\right| < 1$ and have to prove that there exist weights $d = (d_1, ..., d_n, ..., d_1, ..., d_n)$, $d_i > 0$, $\sum_i d_i = 1$ such that $\sum_h \sum_j d_j \left|a^h_{ij}\right| < d_i \ \forall i, \forall h$.

Let us take $d_i = S \max_{h,j} \left|a^h_{ij}\right| + \frac{\overbrace{1 - S \sum_i \max_{h,j} \left|a^h_{ij}\right|}^{>0}}{n} \ \forall i$. These can be taken as weights since $\sum_{i=1}^n d_i = 1, 0 < d_i < 1$.

Notice that $d_i > S \max_{h,j} \left|a^h_{ij}\right| = \underbrace{\sum_{j=1}^n d_j}_{=1} \underbrace{\sum_{h=1}^S}_{=S} \max_{h,j} \left|a^h_{ij}\right| > \sum_h \sum_j d_j \left|a^h_{ij}\right|, \forall i, \forall h.$

B.6.5 Proof of Proposition 4.10

For the case of $n = 1$, the condition for the alternative definition of D–stability simplifies the requirement for Ω to be stable and for at least one of the following to hold true

$$\left(\frac{\frac{1}{\delta_1}}{1+\frac{1}{\delta_1^2}}\left(-\rho_l A_1\right)+...+\frac{\frac{1}{\delta_S}}{1+\frac{1}{\delta_S^2}}\left(-\rho_l A_S\right)\right)\neq 0,$$

$$\left(\frac{1}{1+\frac{1}{\delta_1^2}}\left(-\rho_l A_1\right)+...+\frac{1}{1+\frac{1}{\delta_S^2}}\left(-\rho_l A_S\right)+1\right)\neq 0 \text{ for all } l=0,1,...,k\ (\rho_0=1).$$

The first "same sign" condition follows directly from the first inequality above. The second condition that follows from the second inequality is proved below.

Necessity: Follows directly from the proof of Proposition 4.12. Just note that in the univariate economy setup any sum of minors M_k consists of elements $\delta_{h_1}\delta_{h_2}...\delta_{h_k}(-\rho_l A_{h_1}-\rho_l A_{h_2}-...-\rho_l A_{h_k}+1)$ and that if the sum of nonnegative elements is strictly greater than zero, then at least one of them has to be strictly positive.

Sufficiency: I have $-\rho_l A_{h_1}-\rho_l A_{h_2}-...-\rho_l A_{h_p}+1\geq 0$ for any subeconomy $(h_1,...,h_p)$ and for each group of subeconomies of size p, $\exists\ h_1^*(l),...,h_p^*(l)\vdots -\rho_l A_{h_1^*}-\rho_l A_{h_2^*}-...-\rho_l A_{h_p^*}+1>0$, and have to prove that $\left(\frac{1}{1+\frac{1}{\delta_1^2}}\left(-\rho_l A_1\right)+...+\frac{1}{1+\frac{1}{\delta_S^2}}\left(-\rho_l A_S\right)+1\right)\neq 0$.

I group separately the terms corresponding to non-positive $\rho_l A_i$'s and terms corresponding to strictly positive $\rho_l A_i$'s.

Schematically, I will have

$$\underbrace{\left[\frac{1}{1+\frac{1}{\delta_1^2}}\left(\rho_l A_1^-\right)+...+\frac{1}{1+\frac{1}{\delta_k^2}}\left(\rho_l A_k^-\right)\right]}_{\leq 0}+\underbrace{\left[\frac{1}{1+\frac{1}{\delta_1^2}}\left(\rho_l A_1^+\right)+...+\frac{1}{1+\frac{1}{\delta_m^2}}\left(\rho_l A_m^+\right)\right]}_{\leq 1}-$$

1. If the first sum is strictly less than zero, then the whole expression is less than zero. If the first sum is equal to zero, then the second sum (if there are any positive $\rho_l A'$s at all) has to be less than 1: for the whole economy I have to have that $-\rho_l A_1-\rho_l A_2-...-\rho_l A_S+1>0$, that is, excluding zero $\rho_l A'$s I have to have $-\rho_l A_1^+-...-\rho_l A_m^++1>0$, and also take into account that $0<\frac{1}{1+\frac{1}{\delta_1^2}}<1$, which proves the sufficiency part of the second condition in Proposition 4.10.

B.6.6 Proof of Proposition 4.11

For the case of $n=2$, the inequality in the alternative definition of D–stability looks as follows:

$$\det\left[\frac{-\rho_l A_1}{1+\frac{i}{\delta_1}}+...+\frac{-\rho_l A_S}{1+\frac{i}{\delta_S}}+I\right]=1+\det\frac{(-\rho_l A_1)}{1+\frac{i}{\delta_1}}+...+\det\frac{(-\rho_l A_S)}{1+\frac{i}{\delta_S}}+\frac{M_1(-\rho_l A_1)}{1+\frac{i}{\delta_1}}+...+\frac{M_1(-\rho_l A_S)}{1+\frac{i}{\delta_S}}+$$

$$+\det mix\left(\frac{-\rho_l A_1}{1+\frac{i}{\delta_1}},\frac{-\rho_l A_2}{1+\frac{i}{\delta_2}}\right)+...+\det mix\left(\frac{-\rho_l A_{S-1}}{1+\frac{i}{\delta_{S-1}}},\frac{-\rho_l A_S}{1+\frac{i}{\delta_S}}\right)=$$

$$=1+\left(\frac{1-\frac{i}{\delta_1}}{1+\frac{1}{\delta_1^2}}\right)^2\det(-\rho_l A_1)+...+\left(\frac{1-\frac{i}{\delta_S}}{1+\frac{1}{\delta_S^2}}\right)^2\det(-\rho_l A_S)+$$

$$+\left(\frac{1-\frac{i}{\delta_1}}{1+\frac{1}{\delta_1^2}}\right)M_1(-\rho_l A_1)+...+\left(\frac{1-\frac{i}{\delta_S}}{1+\frac{1}{\delta_S^2}}\right)M_1(-\rho_l A_S)+...+$$

$$+\left(\frac{1-\frac{i}{\delta_1}}{1+\frac{1}{\delta_1^2}}\right)\left(\frac{1-\frac{i}{\delta_2}}{1+\frac{1}{\delta_2^2}}\right)[\det mix(-\rho_l A_1,-\rho_l A_2)+\det mix(-\rho_l A_2,-\rho_l A_1)]+...+$$

$$+\left(\frac{1-\frac{i}{\delta_{S-1}}}{1+\frac{1}{\delta_{S-1}^2}}\right)\left(\frac{1-\frac{i}{\delta_S}}{1+\frac{1}{\delta_S^2}}\right)[\det mix(-\rho_l A_{S-1},-\rho_l A_S)+\det mix(-\rho_l A_S,-\rho_l A_{S-1})]\neq 0$$

for all $l=0,1,...,k,\ (\rho_0=1)$.

Taking real and imaginary parts, one gets

$$\operatorname{Re}\det\left[\frac{-\rho_l A_1}{1+\frac{i}{\delta_1}}+...+\frac{-\rho_l A_S}{1+\frac{i}{\delta_S}}+I\right]=1+\frac{1-\frac{1}{\delta_1^2}}{\left(1+\frac{1}{\delta_1^2}\right)^2}\det(-\rho_l A_1)+...+$$

$$+\frac{1-\frac{1}{\delta_S^2}}{\left(1+\frac{1}{\delta_S^2}\right)^2}\det(-\rho_l A_S)+\frac{1}{1+\frac{1}{\delta_1^2}}M_1(-\rho_l A_1)+...+\frac{1}{1+\frac{1}{\delta_S^2}}M_1(-\rho_l A_S)+...+$$

$$+\frac{1-\frac{1}{\delta_1\delta_2}}{\left(1+\frac{1}{\delta_1^2}\right)\left(1+\frac{1}{\delta_2^2}\right)}[\det mix(-\rho_l A_1,-\rho_l A_2)+\det mix(-\rho_l A_2,-\rho_l A_1)]+...+$$

$$+\frac{1-\frac{1}{\delta_{S-1}\delta_S}}{\left(1+\frac{1}{\delta_{S-1}^2}\right)\left(1+\frac{1}{\delta_S^2}\right)}[\det mix(-\rho_l A_{S-1},-\rho_l A_S)+\det mix(-\rho_l A_S,-\rho_l A_{S-1})]$$

$$\operatorname{Im}\det\left[\frac{-\rho_l A_1}{1+\frac{i}{\delta_1}}+...+\frac{-\rho_l A_S}{1+\frac{i}{\delta_S}}+I\right]=\frac{-\frac{2i}{\delta_1}}{\left(1+\frac{1}{\delta_1^2}\right)^2}\det(-\rho_l A_1)+...+$$

$$+\frac{-\frac{2i}{\delta_S}}{\left(1+\frac{1}{\delta_S^2}\right)^2}\det(-\rho_l A_S)+\frac{-\frac{i}{\delta_1}}{1+\frac{1}{\delta_1^2}}M_1(-\rho_l A_1)+...+\frac{-\frac{i}{\delta_S}}{1+\frac{1}{\delta_S^2}}M_1(-\rho_l A_S)+...+$$

$$+\frac{-i\left(\frac{1}{\delta_1}+\frac{1}{\delta_2}\right)}{\left(1+\frac{1}{\delta_1^2}\right)\left(1+\frac{1}{\delta_2^2}\right)}[\det mix(-\rho_l A_1,-\rho_l A_2)+\det mix(-\rho_l A_2,-\rho_l A_1)]+...+$$

$$+\frac{-i\left(\frac{1}{\delta_{S-1}}+\frac{1}{\delta_S}\right)}{\left(1+\frac{1}{\delta_{S-1}^2}\right)\left(1+\frac{1}{\delta_S^2}\right)}[\det mix(-\rho_l A_{S-1},-\rho_l A_S)+\det mix(-\rho_l A_S,-\rho_l A_{S-1})]$$

for all $l=0,1,...,k$ $(\rho_0=1)$ for all $l=0,1,...,k$ $(\rho_0=1)$.

From the Im part of the determinant I see the "same sign" sufficient condition for this case:

$\det(-\rho_l A_i) \geq 0$, $[\det mix(-\rho_l A_i, -\rho_l A_j) + \det mix(-\rho_l A_j, -\rho_l A_i)] \geq 0$, $i \neq j$, $M_1(-\rho_l A_i) \geq 0$,

or

$\det(-\rho_l A_i) \leq 0$, $[\det mix(-\rho_l A_i, -\rho_l A_j) + \det mix(-\rho_l A_j, -\rho_l A_i)] \leq 0$, $i \neq j$, $M_1(-\rho_l A_i) \leq 0$ for all $l = 0, 1, ..., k$ $(\rho_0 = 1)$

If all inequalities above are equalities to zero, then the real part equals 1, and the sufficient condition for δ–stability holds true.

B.6.7 Proof of Propositions 4.12 and 4.13

I consider $\Gamma = D(-\Omega)$. A necessary and sufficient condition for stability of this matrix is that real parts of eigenvalues of $D(-\Omega)$ be greater than zero. And for the condition on eigenvalues to hold true it is necessary that all sums of principal minors of $D(-\Omega)$ grouped by the same size be greater than zero.

Indeed, the characteristic equation for eigenvalues of Γ has the form

$\det(\Gamma + I\mu) = \det\Gamma + \mu M_{n-1} + \mu^2 M_{n-2} + ... + \mu^{n-1} M_1 + \mu^n = 0$, where $\lambda = -\mu$ is the eigenvalue of Γ. and M_k is the sum of all principal minors of Γ of size k.

On the other hand, the same characteristic equation can be written in terms of the product decomposition of the polynomial:

$$(\mu + \lambda_1) \cdots (\mu + \lambda_n) = \underbrace{\lambda_1 ... \lambda_n}_{>0} + ... + \mu^{n-2} \underbrace{(\lambda_1 \lambda_2 + ... + \lambda_{n-1} \lambda_n)}_{>0} + \mu^{n-1} \underbrace{(\lambda_1 + ... + \lambda_n)}_{>0} + \mu^n = 0.$$

Thus, all $M_k > 0$.

By writing this condition in terms of $D(-\Omega)$, one gets that in each size group the sum of minors is subdivided into groups of sums of minors that contain the same number of columns of each block of $(-\Omega)$, i.e. $A_i - I$. The coefficient before such particular sum has the form $(\delta_{h_1})^{j_1} (\delta_{h_2})^{j_2} ... (\delta_{h_p})^{j_p}$. This coefficient uniquely specifies the sum of minors by the size, the number of columns from each block, and from which subeconomy it is formed, $(h_1, ..., h_p)$. The size of minors in such a group is equal to the total power of the coefficients, $j_1 + ... + j_p$, and the subscripts of δ's denote from which block of $(-\Omega)$ columns are taken, while the power of each δ indicates how many columns are taken from this particular block.

Let us fix one subeconomy (say, formed by blocks 1, 2, 3) and consider the limit of inequalities for the sum of minors, with δ's for other blocks going to zero. Doing the same operation for all subeconomies, I will get condition (*). The statement in Proposition 13 is derived by setting all δ's for all subeconomies in condition (*) equal to 1.

Appendix C

Appendix to Chapter 5

C.1 Proofs of propositions in Chapter 5

C.1.1 Proof of Propositions 5.1 and 5.2

The PLM in general form is $y_t = a + \Gamma w_t$. If w_i is not included in the PLM, it is reflected in the corresponding zero column of Γ. The REE conditions can be written as $(\rho_i A - I_n)\begin{bmatrix} \gamma_{1i} & \dots & \gamma_{ni} \end{bmatrix}' + B^i = 0$, $i \in I_0$.

It is clear that in case i is not included into the active factors set, that is $\begin{bmatrix} \gamma_{1i} & \dots & \gamma_{ni} \end{bmatrix}' = 0$, then in order to have a REE solution, B^i has to be equal to 0, so that one can omit only those factors in the PLM , that have a zero column in B in the reduced form. Equivalently, it is clear that if $B_i \neq 0$, then, in order to have a REE solution, one should not have$\begin{bmatrix} \gamma_{1i} & \dots & \gamma_{ni} \end{bmatrix}' = 0$, that is, one has to include w_i into the active factors set.

In case i is included in the active factors set, that is $\begin{bmatrix} \gamma_{1i} & \dots & \gamma_{ni} \end{bmatrix}' \neq 0$, the REE solution exists if and only if the following conditions holds true.

$B^i = 0$, or ($B^i \neq 0$ and $\det(\rho_i A - I) \neq 0$), or ($B^i \neq 0$ and $\det(\rho_i A - I) = 0$ and $rank(\rho_i A - I) = rank(\rho_i A - I, B^i)$).

Combining the two cases we get the statement in Proposition 5.1.

For Proposition 5.2, one has only to transform the last conditions to guarantee the uniqueness of the solution.

In case i is included in the active factors set, that is $\begin{bmatrix} \gamma_{1i} & \dots & \gamma_{ni} \end{bmatrix}' \neq 0$, the REE solution exists and is unique if and only if the following condition holds true.

$\det(\rho_i A - I) \neq 0$.

C.1.2 Proof of Proposition 5.5 (Necessary conditions and sufficient conditions in terms of eigenvalues for the structurally homogeneous case)

We have to study matrix $D_1 \Omega_{\rho_l}$ for stability under any $\delta_h > 0$, where D_1 and Ω_{ρ_l} are defined in (5.14) and (5.16), respectively. Thus we consider

$$\begin{aligned} \det\left(\Omega_{\rho_l} - D_1^{-1}\mu I\right) &= \det \begin{bmatrix} \rho_l A_1 - \left(1 + \frac{\mu}{\delta_1}\right) I & \cdots & \rho_l A_S \\ \vdots & \ddots & \vdots \\ \rho_l A_1 & \cdots & \rho_l A_S - \left(1 + \frac{\mu}{\delta_S}\right) I \end{bmatrix} = 0, \\ \forall l &= 0, ..., k, (\rho_0 = 1), \end{aligned}$$

where $A_h = \zeta_h A$, $\sum \zeta_h = 1$.

It is clear from the structure of the matrix above that $\mu = -\delta_{i_0}$ is a root if and only if at least one of the following holds true: $\bar{A}$ is singular or there exists at least one other δ_j that equals δ_{i_0}. (If A is singular, then $\mu_h = -\delta_h$, $h = \overline{1, S}$ are the roots. That is, if none of $-\delta'$s is the root, then A is non-singular.)

Assume that A is non-singular and all δ_h's are different, that is assume that none of $-\delta'$s is the root. If there are roots other than $-\delta_h'$s (the case of eigenvalues $\mu_h = -\delta_h < 0$ is obvious), then they satisfy the characteristic equation for obtaining the eigenvalues of $D_1 \Omega_{\rho_l}$ that are not equal to $-\delta_h$:

$$\det\left(\Omega_{\rho_l} - D_1^{-1}\mu I\right) = \det\begin{bmatrix} \rho_l A_1 - \left(1 + \frac{\mu}{\delta_1}\right) I & \cdots & \rho_l A_S \\ \vdots & \ddots & \vdots \\ \rho_l A_1 & \cdots & \rho_l A_S - \left(1 + \frac{\mu}{\delta_S}\right) I \end{bmatrix} =$$

(subtracting the last row from other rows)

$$= \det\begin{bmatrix} -\left(1 + \frac{\mu}{\delta_1}\right) I & 0 & \cdots & 0 & \left(1 + \frac{\mu}{\delta_S}\right) I \\ 0 & -\left(1 + \frac{\mu}{\delta_2}\right) I & \cdots & 0 & \left(1 + \frac{\mu}{\delta_S}\right) I \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\left(1 + \frac{\mu}{\delta_{S-1}}\right) I & \left(1 + \frac{\mu}{\delta_S}\right) I \\ \rho_l A_1 & \rho_l A_2 & \cdots & \rho_l A_{S-1} & \rho_l A_S - \left(1 + \frac{\mu}{\delta_S}\right) I \end{bmatrix} =$$

(for $\mu \neq \delta_h \ \forall h$)

$$= \left(1 + \frac{\mu}{\delta_1}\right) \times ... \times \left(1 + \frac{\mu}{\delta_S}\right) \det\begin{bmatrix} -I & \cdots & 0 & I \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & -I & I \\ \frac{\rho_l A_1}{\left(1+\frac{\mu}{\delta_1}\right)} & \cdots & \frac{\rho_l A_{S-1}}{\left(1+\frac{\mu}{\delta_{S-1}}\right)} & \frac{\rho_l A_S}{\left(1+\frac{\mu}{\delta_S}\right)} - I \end{bmatrix} =$$

(adding all columns to the last one)

$$= \left(1 + \frac{\mu}{\delta_1}\right) \times ... \times \left(1 + \frac{\mu}{\delta_S}\right) \det\begin{bmatrix} -I & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & -I & 0 \\ \frac{\rho_l A_1}{\left(1+\frac{\mu}{\delta_1}\right)} & \cdots & \frac{\rho_l A_{S-1}}{\left(1+\frac{\mu}{\delta_{S-1}}\right)} & \left[\frac{\rho_l A_1}{1+\frac{\mu}{\delta_1}} + ... + \frac{\rho_l A_S}{1+\frac{\mu}{\delta_S}} - I\right] \end{bmatrix} =$$

$$= \left(1+\frac{\mu}{\delta_1}\right)\times ... \times\left(1+\frac{\mu}{\delta_S}\right)(-1)^{n(S-1)}\det\left[\frac{\rho_l A_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\rho_l A_S}{1+\frac{\mu}{\delta_S}}-I\right]=0.$$

As we consider $\mu \neq -\delta_h$, the last equation is equivalent to

$$\det\left[\frac{-\rho_l A_1}{1+\frac{\mu}{\delta_1}}+...+\frac{-\rho_l A_S}{1+\frac{\mu}{\delta_S}}+I\right]=0, \text{ where } A_h=\zeta_h A, \sum\zeta_h=1.$$

After some calculations, we obtain

$$\det\left[\rho_l A\left(\frac{-\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{-\zeta_S}{1+\frac{\mu}{\delta_S}}\right)+I\right]=0,$$

and finally

$$\rho_l\lambda_k\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\zeta_S}{1+\frac{\mu}{\delta_S}}\right)=1$$

for those λ_k, eigenvalues of A, that are not equal to zero. If all $\lambda_k=0$, then A is a zero matrix and the only eigenvalues of $D\Omega$ are $-\delta_h$'s.

As complex eigenvalues of a real matrix A come in conjugate pairs, the system above is equivalent to

$$\begin{cases} \rho_l \operatorname{Re}(\lambda_k)\operatorname{Re}\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\zeta_S}{1+\frac{\mu}{\delta_S}}\right)-\rho_l\operatorname{Im}(\lambda_k)\operatorname{Im}\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\zeta_S}{1+\frac{\mu}{\delta_S}}\right)=1 \\ \rho_l \operatorname{Im}(\lambda_k)\operatorname{Re}\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\zeta_S}{1+\frac{\mu}{\delta_S}}\right)+\rho_l\operatorname{Re}(\lambda_k)\operatorname{Im}\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\zeta_S}{1+\frac{\mu}{\delta_S}}\right)=0 \end{cases}$$

for each pair of conjugate eigenvalues. In case of a real eigenvalue, $\operatorname{Im}(\lambda_k)=0$, the corresponding system simplifies to

$$\rho_l\operatorname{Re}(\lambda_k)\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\zeta_S}{1+\frac{\mu}{\delta_S}}\right)=\rho_l\lambda_k\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+...+\frac{\zeta_S}{1+\frac{\mu}{\delta_S}}\right)=1$$

For any S we have that for eigenvalues μ to be negative, it is necessary that $\frac{\frac{1}{\rho_l\lambda_k}-1}{\frac{1}{\rho_l\lambda_k\delta_1...\delta_S}}>0$ and therefore that $\rho_l\lambda_k<1, \forall l=0,...,k, (\rho_0=1)$. As $|\rho_l|<1, \forall l=\overline{1,k}$, the latter condition is equivalent to $\lambda_k<1$.

For $S=2$, the system corresponding to a real eigenvalue looks as follows:

$$\begin{cases} \rho_l\lambda_k\left(\frac{\zeta_1}{1+\frac{\mu}{\delta_1}}+\frac{\zeta_2}{1+\frac{\mu}{\delta_2}}\right)=1 \\ \mu^2+\mu\frac{\frac{1}{\rho_l\lambda_k}\left(\frac{1}{\delta_1}+\frac{1}{\delta_2}\right)-\left(\frac{\zeta_1}{\delta_2}+\frac{\zeta_2}{\delta_1}\right)}{\frac{1}{\rho_l\lambda_k\delta_1\delta_2}}+\frac{\frac{1}{\rho_l\lambda_k}-1}{\frac{1}{\rho_l\lambda_k\delta_1\delta_2}}=0. \end{cases}$$

The Routh–Hurwitz conditions for negativity of real parts of μ are necessary and sufficient and look as follows:

$$\begin{cases} \frac{\frac{1}{\rho_l\lambda_k}-1}{\frac{1}{\rho_l\lambda_k\delta_1\delta_2}}>0 \\ \frac{\frac{1}{\rho_l\lambda_k}\left(\frac{1}{\delta_1}+\frac{1}{\delta_2}\right)-\left(\frac{\zeta_1}{\delta_2}+\frac{\zeta_2}{\delta_1}\right)}{\frac{1}{\rho_l\lambda_k\delta_1\delta_2}}>0 \end{cases}.$$

The system of inequalities above is equivalent to

$$\begin{cases} \rho_l \lambda_k < 1 \\ \rho_l \lambda_k < \frac{\frac{1}{\delta_1}+\frac{1}{\delta_2}}{\frac{\zeta_1}{\delta_2}+\frac{\zeta_2}{\delta_1}}. \end{cases}.$$

Since $\frac{\frac{1}{\delta_1}+\frac{1}{\delta_2}}{\frac{\zeta_1}{\delta_2}+\frac{\zeta_2}{\delta_1}} > 1$, as $\frac{1-\zeta_1}{\delta_1} + \frac{1-\zeta_2}{\delta_2} > 0$, the last system of inequalities is equivalent to $\rho_l \lambda_k < 1, \forall l = 0, ..., k, (\rho_0 = 1)$. As $|\rho_l| < 1, \forall l = \overline{1, k}$, the latter condition is equivalent to $\lambda_k < 1$.

Thus, we get the sufficient condition for stability for the case of $S = 2$, that all eigenvalues of A are real and less than 1; and the necessary condition for stability for any S is that all real eigenvalues of A have to be less than 1. $Q.E.D.$