Endogenous fluctuations in game theory and macroeconomics



Marco Pangallo
St Cross College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2019

Endogenous fluctuations in game theory and macroeconomics

Marco Pangallo

St Cross College
University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Trinity 2019

At all scales, economies undergo substantial fluctuations instead of changing smoothly over time. There exist two alternative views about these fluctuations. According to the first view, fluctuations are caused by exogenous noise. Absent noise, economies would return to stable stationary states. The second view argues instead that such stationary states are unstable, and economies endogenously fluctuate following limit cycles or chaotic attractors. Endogenous fluctuations would occur even without noise. It is not yet clear which of these two interpretations most closely describes economic reality, but it has fundamental intellectual and policy implications. This thesis presents theoretical arguments in favor of the second view, in both micro- and macro-economic systems.

In the first part, we study the smallest scale at which these two alternative views can be compared. This is a normal form game repeatedly played by two players who learn from their mistakes and form beliefs about their opponents' future play. Under the exogenous view, players would converge to a Nash equilibrium, and strategic fluctuations would only be caused by behavioral noise. The endogenous view argues instead that strategic dynamics may fail to converge to equilibrium. Here, we completely characterize convergence in games with two players and two actions per player under a learning rule known as Experience-Weighted Attraction. We also introduce a theory that we call "best-reply structure" to approximately characterize convergence in games with two players and an arbitrary number of actions. Our results show that non-convergence and so endogenous fluctuations are more common in games than what might be thought.

The second part of this thesis studies the largest scale of economic fluctuations, namely business cycles in macroeconomic systems. The two views propose different mechanisms to explain correlation, or comovement, of business cycles across economic sectors and across countries. The first view argues that comovement is caused by the propagation of shocks exogenous to the economy, such as political decisions and natural catastrophes. According to the second view, comovement is instead at least partially caused by the synchronization of non-linear dynamics that describe the business cycle. Here, we introduce a demand-driven model of endogenous business cycles in which synchronization occurs through input-output linkages. We develop a theory that mathematically illustrates the interplay between synchronizing non-linear dynamics, shock propagation and network structure. While exogenous business cycle models have difficulty to generate a level of comovement that is as high as in the data, synchronization of endogenous business cycles generates much stronger comovement, potentially solving long-standing puzzles in macroeconomics.

Acknowledgements

First and foremost, I would like to thank my supervisor Doyne Farmer. His supervision in these four years has shaped much of the scientist I have become. I really share his research principles of asking important questions without being constrained by disciplinary boundaries and by reliance on specific methods, and of combining interdisciplinary research with domain knowledge. I am also grateful for his guidance on all aspects of academic life: scientific rigor, writing, giving talks, navigating the publication process. Finally, I am thankful for his relaxed, sympathetic and friendly approach, and for the research freedom that I could enjoy.

My DPhil would not have been the same without my INET Oxford colleagues. I learned a lot from all of them, in talks, group meetings, journal clubs and casual conversations at the pub. I especially thank Torsten Heinrich, François Lafond, Penny Mealy, R. Maria del Rio Chanona, and Anton Pichler. I also thank Luca Mungo, Samuel Wiese and Yoojin Yang for their excellent work during their summer internships that extended results presented in Chapter 3 of this thesis.

I would like to thank Tobias Galla for providing detailed comments on the contents of Chapter 2 of this thesis. I also thank the other academics that gave useful feedback on the topics in this thesis, especially Paul Beaudry, Vincent Crawford, Ido Erev, Cars Hommes, Thomas Peron, and Peyton Young.

I am really grateful to INET Oxford, EPSRC (award number 1657725) and Baillie Gifford for financial support.

Thanks to my Oxford housemates Michael, Luca, Hamish, Joe, Federico and Callum for our endless conversations, cooking nights and chocolate cult.

Last but not least, I would like to thank my mother Marzia, my father Enrico and my brother Carlo for their love and emotional support throughout my life, and for their practical support during the DPhil years, including the homemade food supply. Thanks for encouraging me to choose my own path with my brain, but also with my heart.

This thesis is dedicated to Lucrezia, who has always been close by, even when I was in Oxford and she was in Torino, making every day of these last four years happier.

Contents

Introduction **1**

I Endogenous fluctuations in games **13**

1 Literature review **15**

1.1 Overview 15

1.2 Concepts in game theory 16
1.2.1 Notation 16
1.2.2 Nash Equilibrium 17
1.2.3 Classes of normal form games 18

1.3 Convergence properties of some learning algorithms 20
1.3.1 Best reply dynamics 20
1.3.2 Reinforcement learning 22
1.3.3 Fictitious play 25
1.3.4 Replicator dynamics 28
1.3.5 Experience-Weighted Attraction 30
1.3.6 Level-k learning 32
1.3.7 Other learning algorithms 34

1.4 Statistical properties of games 36

2 A taxonomy of learning dynamics in 2x2 games **39**

2.1 Overview 39

2.2 Preliminaries 40
2.2.1 Classes of 2×2 games 40
2.2.2 Special cases of EWA 43
2.2.3 Plan for the exploration of the parameter space 45
2.2.4 Summary of the results 47

2.3 Baseline scenario . . . 50
2.3.1 Simplifications . . . 51
2.3.2 Fixed point analysis . . . 52
2.3.3 Simulations of unstable dynamics . . . 59

2.4 Extensions . . . 61
2.4.1 Unconstrained payoffs . . . 61
2.4.2 Belief learning . . . 62
2.4.3 Reinforcement learning . . . 63
2.4.4 Stochastic learning . . . 65
2.4.5 Other parameter combinations . . . 68

2.A Appendix . . . 70
2.A.1 Replicator dynamics with finite memory . . . 70
2.A.2 Pure strategy fixed points stability . . . 72
2.A.3 Mixed strategy fixed points stability . . . 73
2.A.4 Fixed points of reinforcement learning . . . 74

3 Best reply structure and equilibrium convergence in generic games 77

3.1 Overview . . . 77

3.2 Why should we study randomly-generated games? . . . 78

3.3 How the best reply structure shapes learning dynamics . . . 80

3.4 Quantitative evidence on the effect of the best reply structure . . . 83
3.4.1 Main results . . . 84
3.4.2 Mixed equilibria and the best reply structure . . . 88

3.5 Variation of the best reply structure . . . 92

3.6 Analytical approach . . . 95
3.6.1 Frequency of best reply vectors . . . 97
3.6.2 Other combinatorial calculations . . . 102

3.7 Perspectives . . . 104

3.A Details of the simulation protocol . . . 107
3.A.1 Reinforcement learning . . . 107
3.A.2 Fictitious play . . . 109
3.A.3 Replicator dynamics . . . 111
3.A.4 Experience-Weighted Attraction . . . 113
3.A.5 Experience-Weighted Attraction with noise . . . 115

3.A.6 Level-k learning . . . 117
3.A.7 Initialization of the payoff matrices . . . 117

II Endogenous fluctuations in macroeconomics 119

4 Literature review 121

4.1 Overview . . . 121

4.2 Empirical evidence for non-linear dynamics in economic time series . . . 122

4.3 Endogenous business cycle modeling . . . 125
4.3.1 A taxonomy of endogenous business cycle models . . . 126
4.3.2 A generalized framework based on strategic complementarities . . . 129

4.4 Channels for the synchronization of business cycles . . . 134
4.4.1 Which economic interactions can lead to synchronization? . . . 134
4.4.2 An input-output endogenous business cycle model . . . 140

4.5 A crash course in synchronization theory . . . 145
4.5.1 Phase synchronization . . . 146
4.5.2 Complete synchronization . . . 148

4.6 Previous work on phase synchronization of business cycles . . . 149

5 Synchronization of endogenous business cycles 151

5.1 Overview . . . 151

5.2 Dynamics of uncoupled agents . . . 152

5.3 Phase synchronization . . . 155
5.3.1 Abstract interaction network . . . 155
5.3.2 Input-output business cycle model . . . 161

5.4 Complete synchronization . . . 169
5.4.1 Abstract interaction network . . . 171
5.4.2 Input-output business cycle model . . . 180

5.5 Extensions . . . 185

Conclusions 189

Bibliography 191

Introduction

At all scales, economies undergo substantial fluctuations instead of changing smoothly over time. When two or a few individuals interact strategically playing a game, experiments show that they often fail to coordinate on immutable strategic outcomes, and instead frequently vary their decisions. At intermediate scales defined by single markets, such as housing or financial markets, prices oscillate wildly over time. At the largest macroeconomic scale, economic activity typically experiences a sequence of booms and busts known as the business cycle. These fluctuations at all scales have profound implications on economic behavior and welfare. For example, price volatility increases uncertainty, and business cycles cause large unemployment swings that hurt the well-being and productivity of workers. Understanding the origins of economic fluctuations is one of the most important questions in economics.

There are mainly two interpretations for economic fluctuations. The currently prevailing view is that, at all scales, economies live in stable steady states and would move smoothly from one state to another. However, economies are persistently hit by exogenous noise that prevents them from changing smoothly. The properties of the noise and the way it propagates can explain economic fluctuations. We will call such a view "exogenous fluctuations". The second interpretation, that is now less accepted but is gaining new traction, is that forces internal, or endogenous, to the economy cause fluctuations. According to this view, steady states, if they exist, are unstable, and economic fluctuations are best described by the resulting non-linear dynamics: limit cycles or chaos. Of course, noise can be added on top of these non-linear dynamics, but fluctuations would occur regardless. We will call this interpretation "endogenous fluctuations".[1] This thesis provides theoretical arguments in favor of this latter view of fluctuations, at the smallest and largest scales.

The smallest scale of economic interactions is often modeled as a game with two *players*. The players can choose between two or more different *actions*, and receive a *payoff* that

[1] In practice, the distinction between exogenous and endogenous is somewhat arbitrary as it depends on the "boundaries" of the model. Indeed, noise that is exogenous in one model could be "endogenized" in a more general model. Philosophically, the only truly exogenous noise may be quantum uncertainty. In this thesis, we mathematically define endogenous and exogenous depending on the specific model that is being used: something endogenous is explained by the model (e.g. by the non-linear dynamics implied by the equations of the model), something exogenous is not (e.g. it is a stochastic forcing term).

depends both on their action and on the action of their opponent. To compare the two views of fluctuations, suppose that the players play the game repeatedly, learn from their success or failure to realize satisfactory payoffs, and form beliefs about the actions that their opponent may play. Over time, players may start playing the same action at every round of the game: They may have converged to what is known as a pure strategy Nash equilibrium. Occasionally, one of the players may randomly experiment with another action, but then she would return to equilibrium. This narrative corresponds to exogenous fluctuations in games. If instead players endlessly change the action they play, either following fairly regular cycles or highly irregular action sequences, it means that their learning dynamics have failed to converge to the equilibrium. Non-convergence is due to certain properties of the game and of the learning algorithm—fluctuations are endogenous.

More generally, given that some games do not have pure strategy Nash equilibria, players' *probabilities* to play actions, also known as *mixed strategies*, may converge to a fixed point, known as mixed strategy Nash equilibrium. In this equilibrium, players' action sequences, sampled from mixed strategies, are i.i.d. If, due to random sampling, a player chooses a certain action for several consecutive times, all players may change their mixed strategies, and start playing other actions more often. However, if the mixed strategy Nash equilibrium is stable under learning, these deviations from i.i.d. play would only be momentarily, and players would return to equilibrium. If instead the mixed strategy Nash equilibrium is unstable, players' probabilities may endogenously fluctuate and never settle to a fixed point.

These two alternative outcomes are explicitly shown and compared to real data in the left panels of Figure 1. The top left panel shows data from an experiment (Suppes and Atkinson, 1960; Erev and Roth, 1998) in which two human subjects repeatedly played a game in which they could choose between two possible actions. In this game, for any combination of actions, the sum of the two payoffs to the two players is constant. This constant-sum game does not have any pure strategy Nash equilibrium. Therefore, under the exogenous fluctuations view, players' mixed strategies should converge to the unique mixed strategy Nash equilibrium, and play should be i.i.d. This is exactly what happens in the middle left panel, in which we simulate play under a learning algorithm known as Experience-Weighted Attraction (Camerer and Ho, 1999). In this panel, we choose the behavioral parameters in a way that the mixed strategies converges to the stable equilibrium. In the bottom left panel, we instead choose the parameters in a way that the mixed strategy dynamics converges to a cycle. This cycle is not perfectly periodic because of sampling noise, but play is far from i.i.d. For this specific example, endogenous fluctuations seem more in line with the data, as experimental play is far from i.i.d., too.

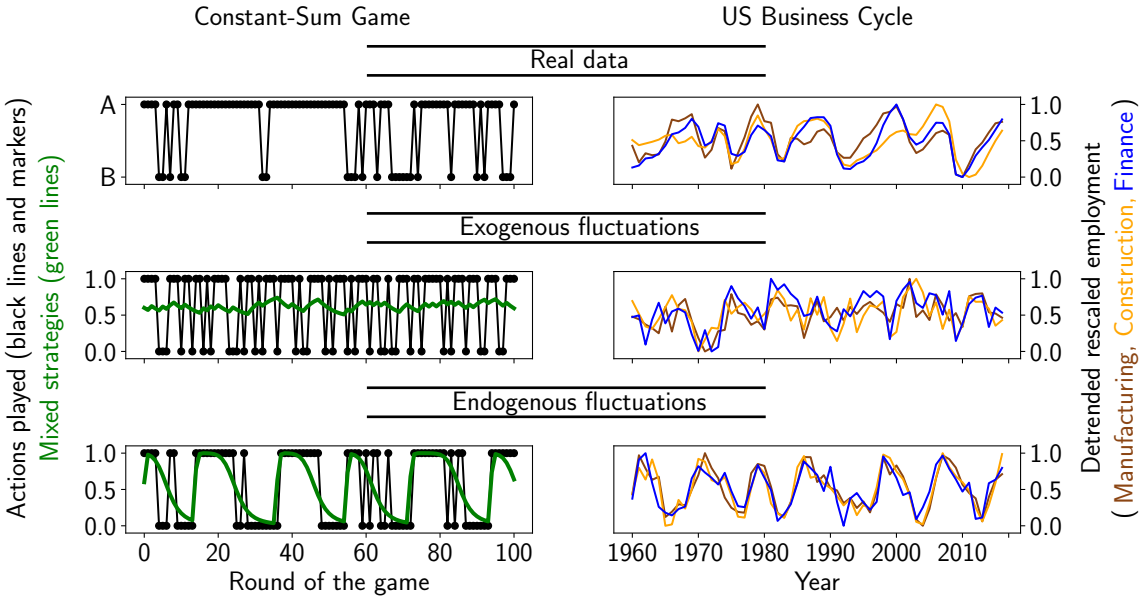


Figure 1: Graphical abstract of this thesis. In the left panels, we show fluctuations in games. The *top left panel* shows data from an experiment in which two individuals, 1 and 2, repeatedly played a constant-sum game with two possible actions, A and B. Markers indicate the actual choice of player 1 at each round of the game; black lines are drawn to guide the eye. The *middle left panel* shows results from the simulation of a learning algorithm that converges to equilibrium. The green line is the mixed strategy of simulated player 1, that is her probability to play action A. Markers are the actions played by the same simulated player, randomly sampled from her mixed strategy. The *bottom left panel* is analogous, but in this case the learning algorithm does not converge to equilibrium, as can be seen by the mixed strategy following an irregular cycle. The right panels show fluctuations in macroeconomics. In the *top right panel*, we show detrended employment in the United States in three economic sectors (manufacturing, construction, finance), rescaled between 0 and 1, from 1960 to 2016. The *middle right panel* shows fluctuations generated by the model presented in this thesis, parameterized so that simulated sectors would converge to a steady state if they were not hit by exogenous shocks. The *bottom right panel* is analogous, but the model is parameterized so that simulated sectors follow a stochastic limit cycle and synchronize their non-linear dynamics.

In this thesis, we first analyze the theoretical behavior of players in all games with two players and two actions per player, under Experience-Weighted Attraction. We show that endogenous fluctuations are only possible in games with no pure strategy Nash equilibria. An example of these games is Matching Pennies: the two players, Alice and Bob, can choose between the two actions Heads and Tails. Alice wins payoff 1 when they both choose the same action (e.g. Heads-Heads), Bob wins the same payoff when they choose different actions (e.g. Heads-Tails). The loser gets payoff -1 in all cases. This game has a *best reply cycle*: starting from both Alice and Bob playing Heads, Bob may switch to Tails (and so win 1, if Alice does not switch), then Alice may switch to Tail, too, and so on. We develop a theory that approximately quantifies the frequency of equilibrium convergence of a family of learning algorithms based on how many best reply cycles the game has. In games with two players, we show that, as the number of actions and the level of "competition" of the game increase, best reply cycles become ubiquitous. This makes convergence to equilibrium

very difficult, and so endogenous fluctuations likely, in many games.

Our analysis of learning dynamics is only theoretical for several reasons. First, because only actions are generally observable in experiments with human subjects, it is difficult to track the evolution of mixed strategies, and so convergence can only be assessed indirectly. Moreover, in experiments there is considerable noise and heterogeneity across players, and it is very difficult to describe the behavior of any player by a single, fixed, learning algorithm. Previous work in experimental game theory averaged play across time to get rid of noise (Erev and Roth, 1998), but this could identify learning dynamics as convergent even when they are not. An experimental assessment of the convergence of learning would need very carefully designed experiments (Cason et al., 2013), which we do pursue here.

The right panels of Figure 1 are concerned with macroeconomic fluctuations. In macroeconomics, the exogenous fluctuations view argues that business cycles are caused by unpredictable events, or *shocks*, such as political decisions, natural catastrophes or wars. Households and firms react to these shocks and propagate them within the economy (Frisch, 1933; Slutzky, 1937). If shocks did not occur, economies would grow smoothly over time. The endogenous fluctuations view claims instead that macroeconomic systems are fundamentally unstable, for example because debt relentlessly builds up to the point that it becomes unsustainable (Minsky, 1992). Therefore, booms sow the seeds for subsequent busts. Negative shocks can still trigger a recession, but they are much more likely to do so if the economy has been in a boom for a while, rather than having just started a recovery from a recession.

The two views make different predictions regarding at least two fundamental properties of macroeconomic time series. The first property is periodicity, exemplified by the existence of peaks in the spectral density of the time series. Under exogenous fluctuations, there should not be any detectable periodicity and so any peak in spectral density. Under endogenous fluctuations, in the specific case in which the business cycle is modeled as a stochastic limit cycle, periodicity would not be perfect, but it should be possible to detect peaks in spectral density. The second property is the correlation, or comovement, across economic sectors or countries. The exogenous fluctuations view attributes comovement to the propagation of shocks only, whereas the endogenous view combines shock propagation with the synchronization[2] of the non-linear dynamics that describe the business cycle. Because of this additional effect, comovement can be expected to be stronger under the endogenous fluctuations view.

[2]The word "synchronization" is often used casually in economics as a synonym of comovement. In this thesis, we use the word synchronization in the dynamical systems sense. In a system made by two or more interacting components, synchronization is the alignment of non-linear dynamics of the components.

In the top right panel of Figure 1, we show employment time series across three large sectors of the economy of the United States: manufacturing, construction and finance.[3] To better distinguish short-term fluctuations from long-term structural change—for example, the secular decline of the employment share of manufacturing—we detrend the data using a filter that removes all fluctuations with periodicities longer than 25 years (Christiano and Fitzgerald, 2003). We further rescale the data so that the minimum of each time series corresponds to zero and the maximum corresponds to one. The data show both a detectable 10-year periodicity, confirming recent results by Beaudry et al. (2019), and strong comovement of employment across sectors. In other words, during economic booms firms in all sectors hire workers, while during busts all firms fire workers.

In the middle right panel, we show simulated data obtained by a business cycle model that we develop in this thesis. In our model, economic sectors are connected via input-output linkages, that is requirements of intermediate goods by a sector to be used as input to produce an output. Input-output linkages are calibrated from real data. We parameterize the model so that simulated sectors would converge to a steady state if the model was deterministic. In line with the exogenous fluctuations view, it is hard to detect any periodicity, and comovement, only caused by shock propagation, is rather weak. In the bottom right panel, we instead parameterize the model so that simulated sectors follow a stochastic limit cycle. We observe higher periodicity and comovement, qualitatively more in line with the data. This intuitively confirms our conjecture above, that comovement is stronger under endogenous fluctuations.[4]

In this thesis, we find the conditions under which sectors with different frequencies of oscillation converge to a common frequency. We show that the power of each sector in determining the common frequency of oscillation depends both on its size and interconnectedness. We then develop a theory that quantifies the balance between the synchronizing effect of input-output linkages and the desynchronizing effect of shocks. Overall, our results provide a mathematical foundation to the potential of endogenous business cycle models to generate comovement.

This is important because comovement across sectors and across countries is typically difficult to obtain in models based on exogenous shocks. Comovement across sectors within a country is necessary to observe aggregate business cycles; a large literature has tried to explain aggregate fluctuations from idiosyncratic shocks hitting economic sectors, but

[3]The data are obtained from the KLEMS database. The sectors correspond to the ISIC 3 codes D, F, and JtK, respectively.

[4]This example is just for illustrative purposes. We plan to undertake a systematic statistical analysis comparing exogenous and endogenous fluctuations in terms of their ability to reproduce empirical time series patterns in future work.

this literature so far has only met partial empirical success. Comovement across national economies is also an important problem, especially for countries that are part of monetary unions (Frankel and Rose, 1998). Yet, international comovement is not much understood, to the point that the inability of models based on trade and exogenous shocks to generate a level of comovement even close to the one in the data has been dubbed "trade-comovement puzzle" (Kose and Yi, 2006). Modeling macroeconomic fluctuations as arising from non-linear dynamics could both address the problems of the microeconomic origin of aggregate fluctuations and of the international comovement.

We now turn to a more detailed description of the approach that we follow and of the main results.

Part I: Endogenous fluctuations in games. The question about whether learning converges to equilibrium is notoriously unresolved in game theory, as the answer seems to depend on the details of the game and the learning algorithm. A large literature has dealt with this problem. Indeed, if it was possible to prove that reasonable learning algorithms converge, this would lend support to the use of equilibrium solutions in game theory. Because equilibrium is a fundamental solution concept, it is unsurprising that convergence of learning algorithms has received a lot of attention.

In Chapter 1, we provide a literature review of the convergence properties of some learning algorithms. Here we study the convergence properties of learning rules that have support from experiments with human players (Camerer, 2003). We do not study whether it is possible to design learning algorithms that have good convergence properties (Nisan et al., 2007). This is a very important field of research at the intersection of mechanism design, engineering and computer science, but here we are interested in whether real players can be expected to converge to equilibrium. The learning algorithms that we consider are reinforcement learning, fictitious play, replicator dynamics, experience-weighted attraction and level-k learning. When they converge to a fixed point, these learning algorithms converge to a Nash equilibrium, or to a point near one. Together with reviewing the literature, in Chapter 1, we also introduce notation, some basic properties of games, and the formal definitions of the learning algorithms. These definitions will be used in the following chapters.

In Chapter 2 we start our analysis of the convergence of learning from the simplest two-player, two-action (2×2) games. These games are simple enough to make it possible to fully characterize the convergence properties of Experience-Weighted Attraction (EWA) in generic 2×2 games. Because EWA generalizes reinforcement learning, fictitious play and replicator dynamics, by studying its convergence properties we provide a taxonomy of

learning dynamics in 2×2 games. This is first of all useful from an experimental point of view. Indeed, EWA is used to model behavior in many experiments, so our analysis provides theoretical guidance on whether learning dynamics can be expected to converge to equilibrium. From a theoretical point of view, we generalize results known in the literature for the learning algorithms that are particular cases of EWA, obtaining interesting new insights. For example, we identify a bifurcation between convergent and unstable learning dynamics in games in the same class as Matching Pennies. This bifurcation depends on two behavioral parameters (memory loss and payoff sensitivity) in a non-trivial way.

We then study generic two-player games with an arbitrary number of actions (see Chapter 3). Because exact results are no longer possible, we develop an approximate indicator of convergence based on a topological property of every normal-form game, that we call best reply structure. We validate our indicator using extensive numerical simulations. Obtaining approximate results and using numerical simulations is unconventional in the game theory literature, which tends to be mathematically rigorous. However, we think it is of high value when one is interested in the convergence properties of arbitrary games that may not have special mathematical properties that make rigorous proofs possible. Also our focus on games with generic payoffs is different from most of the literature. For example, many proofs of convergence apply to potential games (Monderer and Shapley, 1996b), in which all differences in payoffs for unilaterally changing strategy can be expressed using a global potential function.

More precisely, to sample the space of generic games we generate payoffs at random. To understand how studying the typical properties of ensembles of randomly generated systems is useful, we make an analogy to theoretical ecology (see Section 3.2 for a more detailed discussion). In a seminal paper, May (1972) studied randomly-generated ecosystems to understand if they were stable. His key result was that ecosystems tend to become unstable as they grow larger, in contrast with empirical evidence (there exist multiple large ecosystems that are stable within reasonable time frames). Of course May was not assuming that real ecosystems were random. Rather, he was using randomly-generated ecosystems as a null model, that can be sharpened by adding constraints that have empirical support. His work generated a vast body of literature that is still active today (Johnson et al., 2014). In Chapter 3 we take this approach in game theory, by first generating payoffs uniformly at random and then adding constraints that represent deviations from well-known classes of games.

Our indicator of non-convergence is based on best reply dynamics,[5] under which each

[5] Best reply dynamics is well-known in game theory and is widely used to develop intuition about learning, but we use it in a new way, to obtain the approximate probability of convergence in generic games.

player responds with the best reply to his opponent's last action. Best reply dynamics can converge to cycles or fixed points (fixed points of best reply dynamics are pure strategy Nash equilibria). Our indicator is the relative "size" of cycles with respect to fixed points, a quantity that is very simple to compute in every two-player normal-form game. We show that this measure predicts the non-convergence frequency of the learning algorithms that we study in this thesis. We then study how the best reply structure changes when we vary key properties of the ensemble of randomly generated games, such as the number of actions available to the players and the correlation of the payoffs. We find that as games become more complicated and more competitive, in the sense that the number of actions increases and the payoffs become anti-correlated, best reply cycles become dominant, making equilibrium convergence unlikely.

Part II: Endogenous fluctuations in macroeconomics. We start Chapter 4 by asking whether it is possible to assess whether macroeconomic fluctuations are exogenous or endogenous by looking at time series data alone. We discuss the literature that looks for statistical signatures of limit cycles and chaos in the data. The outcome of this strand of research has been somewhat disappointing, in the sense that it has not provided conclusive evidence neither in support nor against one of the two views of fluctuations. Indeed, because of shocks and measurement errors, one would need a lot of observations to support either view, but economic time series are extremely short relative to the typical time scale of fluctuations. We point out that it is necessary to complement time series analysis with theoretical modeling.

After reviewing the main existing endogenous business cycle models, we introduce a general framework based on Beaudry et al. (2019) that generalizes many of the reasons for endogenous fluctuations that have been proposed in the literature. Our framework is based on the idea of *strategic complementarities*, a concept initially formulated within game theory. Supposing that the actions of a game can be given an ordinal meaning, strategic complementarities are the tendency of agents to increase their action if the other agents increase their action, too. In macroeconomics, this can be thought of as the tendency of firms to increase production if other firms increase production themselves, or the tendency of households to increase consumption if other households in their social network increase consumption, too. Strategic complementarities generically lead to unstable steady states; in our framework, they lead to limit cycles, which can be augmented by some shocks to generate realistic imperfectly periodic dynamics. Our framework remains abstract, as we do not model specific causes for fluctuations and specific behaviors of households and firms. However, we view the generality of our framework as an advantage, as it makes it possible

to illustrate the mechanisms for synchronization in a way that does not depend too strongly on the specificity of any given model.

In our framework, synchronization occurs through the main channels identified in the literature. A possible, baseline, channel is common shocks hitting all synchronizing agents. Other channels are demand of intermediate goods by firms to be used in further production, and demand of final goods by consumers to be used for consumption. We model these interaction patterns using the standard concept of input-output network. Input-output linkages exist both across sectors within the same country and across countries. Our framework can thus explain both synchronization of economic activity across sectors—and so the emergence of aggregate fluctuations from sectoral oscillations—and the international synchronization of business cycles.

In Chapter 5, we analyze the synchronization properties of the endogenous business cycle model that we introduced in Chapter 4. We first show that different sectors may follow limit cycles that have different frequencies, if certain parameters characterizing their dynamics are different. We use numerical simulations to assess whether sectors may be able to converge to a common frequency of oscillation, depending on the frequency spread and on the interaction network. Some of our findings confirm well-known results in synchronization theory, but other findings, to the best of our knowledge, are novel. For example, we find that both the size of a sector and its interconnectedness in the input-output network determine the power of that sector to set the common frequency of oscillation. We show that this power is captured by the eigenvector centrality of the sector in a certain normalization of the input-output network. All results discussed so far also apply when the basic unit of analysis is a country, or a sector-country pair, for any number of agents.

After discussing synchronization in the sense of convergence to a common frequency of oscillation ("phase synchronization"), we discuss synchronization in the sense of perfect alignment of non-linear dynamics ("complete synchronization"). Because we study a limit cycle model, if frequencies of oscillation are very close, absent shocks, complete synchronization occurs at all times. However, sectors are hit by idiosyncratic shocks, which act as a desynchronizing force. Here, we extend the master stability function approach (Pecora and Carroll, 1998), which in the original formulation gives a yes/no answer to whether complete synchronization is possible, to quantify *how strong* is coupling of non-linear dynamics with respect to idiosyncratic shocks. We show analytically that this depends on structural properties of the interaction network, captured by the eigenvalues and eigenvectors of the normalized Laplacian matrix of the network.

Previous attempts to introduce synchronization theory into economics have had limited impact. We conjecture that this is due to limitations of the proposed approaches. For ex-

ample, Selover and Jensen (1999) tried to explain international comovement through phase synchronization of a standard model used in synchronization theory (Kuramoto, 1975), that has no explicit economic mechanisms. Other researchers analyzed synchronization of endogenous business cycle models that produced fluctuations with no persistence, unlike in economic time series data, or for simplicity considered at most two sectors or two countries. Our approach can solve these limitations at once, making the potential impact of synchronization theory in economics clearer. Because synchronization theory builds on non-linear dynamics, this would also indirectly support the endogenous view of economic fluctuations.

The big picture. On a high-level, the results of this thesis give support to a complex systems view of economics. It is uncontroversial that the economy is a complex system, as it is formed by billions of heterogeneous individuals that interact with one another at several scales, producing aggregate phenomena that cannot directly be mapped to individual behaviors. Controversies arise on what is the best way to mathematically model the complexity of the economy. According to a broad interpretation, it is sufficient to augment the traditional theoretical framework in economics with heterogeneity, networks, bounded rationality and non-linear dynamics. These have been and currently are productive research agendas in mainstream economics. A more radical, and narrower, interpretation, which we call *complexity economics*, is that the traditional theoretical framework, based on optimization and *equilibrium*, is not well suited to deal with the complexity of the economy.

The word equilibrium means different things in different subfields of economic theory, yet all equilibrium concepts have a common meaning, that in principle has *nothing to do* with stability and stationarity. Indeed, many equilibrium concepts are dynamic and so for example it is possible to have chaotic equilibria.

Equilibrium in economics can generally be thought of as a fixed point in function space, in which beliefs, plans and outcomes are mutually consistent. Economic agents can think, and so have beliefs about states of the economy. Behavioral rules that can be fully or boundedly rational map these beliefs into planned actions. Finally, outcomes resulting from the combined actions of all agents may let each agent realize their planned actions, or may force some agent to choose an action that was not planned. Equilibrium outcomes are such that agents – at least on average – always choose the action that was planned given their beliefs and behavioral rules. In other words, beliefs and plans match outcomes.

We think that assuming equilibrium is a strong top-down constraint imposed on the system. It helps the analysis in several ways, but it is unclear whether it is empirically relevant. Complexity economics models—such as Agent-Based Models—let outcomes emerge

in a bottom-up way without imposing equilibrium constraints, which we think is more in line with a complex systems view of the economy.

The general concept of equilibrium can coincide with stationarity. For example, in game theory players have beliefs about what their opponents will do and plan actions according to these beliefs and some behavioral rule. If players are fully rational, their behavioral rule is to select the action that maximizes their payoff given their beliefs. In a Nash equilibrium all players' actions and beliefs are mutually consistent, so no agent can improve her payoff by switching to another action. The Nash equilibrium is a static equilibrium concept, so equilibrium convergence implies stationarity.

However, it is possible to define more sophisticated equilibrium concepts that are not static. In strategically repeated games, players' strategies are sequences of actions over a certain number of time periods, rather than an action played at any given time period. An equilibrium corresponds to a combination of sequences for all players. The well-known Folk theorem shows that, under some conditions, there exist infinite equilibria, and each equilibrium corresponds to a different combination of sequences. It is then possible to construct equilibria for which the action sequences are cyclic for all players. This is still an equilibrium, because all players believe that their opponents will keep playing the same sequence, behave according to this belief and, in equilibrium, their beliefs correspond to the outcome of the game.

In this thesis, we do not observe these types of equilibria. In our simulations—based on learning algorithms that experiments have shown best describe the behavior of real players—when convergence to (static) Nash equilibrium fails, we often observe chaotic learning dynamics (Skyrms, 1992; Galla and Farmer, 2013). When this happens, the players do not converge to any sort of intertemporal "chaotic equilibrium" (Blume and Easley, 1982; Boldrin and Montrucchio, 1986; Hommes and Sorger, 1998), in the sense that their expectations do not match the outcomes of the game even in a statistical sense.

One could argue that players would be able to recognize patterns in play, and eventually converge to an equilibrium (either static or dynamic). However, in many cases the resulting attractor is high dimensional, making it difficult for a "rational" player to outperform other players by forecasting their actions using statistical methods. Once at least one player systematically deviates from equilibrium, learning and heuristics can outperform equilibrium thinking (Gigerenzer and Todd, 1999) and can be a better description for the behavior of players.

In sum, our results in the first part of this thesis give some theoretical foundation to the idea that players' behavior, at least in some cases, should be modeled out-of-equilibrium.

For the reasons given above, we view this modeling choice more in line with the complex systems nature of the economy.

In the second part of this thesis we focus on macroeconomic systems, for which many dynamic extensions of equilibrium have been put forward. For example, in many macroeconomic models agents are forward-looking, e.g. they plan how much to consume in each future period of their life. These consumption decisions depend, among other things, on future interest rates: in periods when the interest rates are high, agents may prefer saving to consuming, so to earn higher interest and afford higher consumption in the future. In a *rational expectations equilibrium*,[6] the expectations for future interest rates are on average correct, so that again beliefs and planned actions (consumption decisions) match outcomes (interest rates). The assumption of rational expectations places no restriction on macroeconomic dynamics: this may reach a stationary state, but also follow limit cycles or chaotic attractors.

The simple macroeconomic model that is studied in this thesis does not consider rational expectations. It could probably be possible to extend it so that it does, using recently developed methods (Beaudry et al., 2019; Galizia, 2018). An advantage of using rational expectations in an endogenous business cycle model is that fluctuations would occur even if agents were perfectly rational and perfectly able to predict the path of the economy. This shows that the identified source for fluctuations does not depend on agents making systematic mistakes. While we see value in this approach, we think that real individuals do make systematic mistakes, and this is potentially a relevant mechanism to generate endogenous fluctuations, together with other more "fundamental" economic forces. Moreover, we find it clearest to illustrate synchronization dynamics in a model that does not feature rational expectations. This contributes to an increasing body of research suggesting that certain problems in macroeconomics are better dealt in an agent-based modeling setting (Dawid and Delli Gatti, 2018).

Independently of whether one prefers to model economic fluctuations as equilibrium dynamics or not, we believe that the emergence of endogenous oscillations at all scales is a distinctive signature of the complex systems nature of the economy.

[6] In macroeconomics, beliefs over time are called expectations.

Part I

Endogenous fluctuations in games

Chapter 1

Literature review

1.1 Overview

Do players who learn from repeatedly playing a game converge to an equilibrium, for example a Nash equilibrium? Trying to answer this question has been a very active research agenda since the origins of game theory. Indeed, players are most often assumed to play equilibrium, yet questionable assumptions such as common knowledge of rationality make it controversial that they be able to instantly coordinate on one. It is instead universally accepted that players may be able to converge to equilibrium if they learn over repeated play. Under which conditions do they converge?

In general, the answer to this question depends both on the game and on the learning algorithm. The survey presented in this chapter focuses on the literature that is most directly connected to the contributions in the following chapters. It starts with a primer that summarizes the relevant concepts in game theory (Section 1.2). Because in this thesis we focus on normal-form games and on adaptive learning algorithms that have support from experiments with human players, in Section 1.3 we review this subset of the literature. This restriction leaves out the literature on extensive form games and strategically repeated games (e.g. tit for tat). It also excludes the literature on "sophisticated" learning algorithms that are used in mechanism design, engineering and computer science, although for comparison we give a brief account of some of these algorithms in Section 1.3.7.

For the reasons given in the Introduction and in Section 3.2 in more detail, in Chapter 3 we analyze *ensembles* of randomly generated games. We review previous work on ensembles of games in Section 1.4.

More complete books on the theory of learning in games and on equilibrium convergence are Fudenberg and Levine (1998) and Young (2004). Some books with a stronger focus on evolutionary game theory are Vega-Redondo (1996), Hofbauer and Sigmund (1998) and

Gintis (2000). Finally, Camerer (2003) focuses on behavioral game theory and includes one chapter on learning from a behavioral perspective (Chapter 6, pp. 265-335).

1.2 Concepts in game theory

We begin introducing basic concepts in game theory. After explaining the notation, in Section 1.2.2 we define a Nash equilibrium and give a brief account of other equilibrium concepts. Finally, in Section 1.2.3 we define some classes of normal form games that will be used in the following.

1.2.1 Notation

A normal form non-cooperative game is characterized by a set of players, a set of pure strategies (actions) available to each player, and a set of payoffs for each profile (combination) of actions by the players (Myerson, 2013).

In this thesis we focus on two-player games. The two players are indexed by $\mu = R, C$, where R stands by player Row, and C stands by player column ($-\mu$ denotes the other player than player μ). The pure strategies are indexed by $i, j = 1, ..., N$. We denote by i the actions of player Row and by j the actions of player Column. Throughout this thesis, the number of pure strategies is the same for all players. The payoff matrix for player μ is denoted by Π^{μ}, with Π^{μ}_{ij} as the element of the payoff matrix if player μ plays action i and the other player plays action j. So if player Row plays action i and player Column plays action j, they receive payoffs Π^{R}_{ij} and Π^{C}_{ji} respectively. A payoff bi-matrix conveys the information on the payoffs of both players in the same mathematical object. In each cell of the matrix, the number before the comma represents the payoff that player Row receives, whereas the number after the comma is the payoff that player Column receives. For example, in the case of a game with $N = 2$,

$$\Pi = \begin{pmatrix} \Pi^{R}_{11}, \Pi^{C}_{11} & \Pi^{R}_{12}, \Pi^{C}_{21} \\ \Pi^{R}_{21}, \Pi^{C}_{12} & \Pi^{R}_{22}, \Pi^{C}_{22} \end{pmatrix}. \tag{1.1}$$

Instead of being constrained to play a given action with certainty, players can play *mixed strategies*; in this case they select one of their actions with a given probability. Denote the probability that player Row plays action i by x_i and the same quantity for action j and player Column by y_j. (With some abuse of notation, when a player is not specified, x^{μ}_{i} will denote action i by player μ.) Both mixed strategy vectors $\mathbf{x}$ and $\mathbf{y}$ are defined in the (N-1)-dimensional simplex Δ in $\mathbb{R}^N$ (the combination of the two simplices of the two players will be referred to as the "strategy space"):

$$\mathbf{x}, \mathbf{y} \in \Delta = \left\{ \mathbf{p} \in \mathbb{R}^N : p_i \geq 0; \sum_{i=1}^{N} p_i = 1, \forall i = 1, ...N \right\}. \tag{1.2}$$

The pure strategies i and j are particular cases of the mixed strategies, in which one of the actions is chosen with certainty. In mathematical terms, pure strategies are the unit vectors of the canonical basis of $\mathbb{R}^N$. A mixed strategy profile σ is a combination of mixed strategy vectors

$$\sigma = (\mathbf{x}, \mathbf{y}) \in \Delta \times \Delta. \tag{1.3}$$

The payoff to player μ associated with a mixed strategy profile $\sigma = (\mathbf{x}, \mathbf{y})$ is the expected payoff:

$$\Pi^{\mu}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i y_j \Pi_{ij}^{\mu}. \tag{1.4}$$

Note that while we use subscripts as a shorthand to denote payoffs for actions i and j, Π_{ij}^{μ}, here we use parentheses to indicate that the payoff is a function of the mixed strategies. Moreover, we will use often use the shorthand $\sum_{ij}$ to indicate $\sum_{i=1}^{N} \sum_{j=1}^{N}$.

In this thesis, we assume that the normal form game that we described so far is played repeatedly over time, with infinite horizon. Time is indexed by t and the learning dynamics is completely characterized by the components of the mixed strategy vectors $x_i(t), y_j(t), i, j = 1, \ldots, N$. We denote by $s^{\mu}(t)$ the action which is actually taken by player μ at time t, and by $s^{-\mu}(t)$ the action taken by her opponent at the same time. We indicate the payoff in this case by $\Pi^{\mu}(s^{\mu}(t), s^{-\mu}(t))$: this notation, consistently with the choice for mixed strategies, uses parentheses in place of subscripts because $s^{\mu}(t)$ and $s^{-\mu}(t)$ do not denote specific actions. We also use parentheses when the action of one player is specified but the action of the other player is not. For example, if we want to indicate the payoff for player μ playing action i against her opponent playing an unspecified action $s^{-\mu}(t)$, we use $\Pi^{\mu}(i, s^{-\mu}(t))$.

1.2.2 Nash Equilibrium

A profile of (pure or mixed) strategies is a Nash Equilibrium (NE) if no player can unilaterally change her strategy and obtain a higher payoff. A strategy profile $\sigma^{\star} = (\mathbf{x}^{\star}, \mathbf{y}^{\star})$ is a NE if it satisfies:

$$\begin{aligned} &\Pi^{R}(\mathbf{x}^{\star}, \mathbf{y}^{\star}) \geq \Pi^{R}(\mathbf{x}, \mathbf{y}^{\star}), \forall \mathbf{x} \in \Delta, \\ &\Pi^{C}(\mathbf{x}^{\star}, \mathbf{y}^{\star}) \geq \Pi^{C}(\mathbf{x}^{\star}, \mathbf{y}), \forall \mathbf{y} \in \Delta. \end{aligned} \tag{1.5}$$

The famous result by Nash (1950) states that there exists at least one NE in any N-strategy, P-player game, for finite N and P. In practice, we will deal with a multitude of pure and mixed NE.

Nash Equilibria are by no means the only equilibrium concept in game theory, and also not the only equilibrium concept that is used in the literature about the convergence of learning in games. Other common equilibrium concepts in normal-form games[7] are: trembling-hand perfect and proper equilibria (Selten, 1975; Myerson, 1978), refinements of NE that take into account that players can make mistakes; and correlated equilibria (Aumann, 1987), a generalization of NE that assumes that players can react to common signals.[8] In this thesis we focus on a set of learning algorithms that have support from experiments with human players. When they converge to a fixed point, these algorithms converge to a Nash equilibrium (or to a point nearby), so in this thesis we will focus on Nash equilibria.

1.2.3 Classes of normal form games

Most of the literature on the convergence of learning in games states results for particular classes of games with certain mathematical properties. Here we define some of them.

Constant-sum games. The sum of the payoffs to the two players at a given strategy profile is the same for all strategy profiles:

$$\Pi_{ij}^{\mu} + \Pi_{ji}^{-\mu} = m, \forall i, j \in 1, \ldots, N. \tag{1.6}$$

A particular case, with $m = 0$, is *zero-sum games*. A common example is Matching Pennies. In this game, depending on the outcome, one player wins a penny and his opponent loses a penny: the sum of the payoffs is zero. As we will show, typically constant-sum games only have mixed strategy Nash equilibria, however it is possible to construct constant-sum games that have pure strategy Nash equilibria.

Dominance-solvable games. Action i of player μ is *strictly dominated* by action j of the same player if action j yields higher payoff for all actions k of player $-\mu$ (Myerson, 2013):

$$\Pi_{ik}^{\mu} < \Pi_{jk}^{\mu}, \forall k \in 1, \ldots, N. \tag{1.7}$$

If μ is a rational player, she would never choose action i. Iterated elimination of (strictly) dominated actions is the process of removing all actions that are strictly dominated. This process is iterated in the sense that some actions may not be dominated in the initial game,

[7] Many more equilibrium concepts exist in extensive-form games and repeated games.

[8] This equilibrium concept is particularly popular in a recent stream of the literature on the convergence of learning in games. This is because finding correlated equilibria is less computationally expensive than finding NE (Nisan et al., 2007), and so some learning algorithms can converge to correlated equilibria in all games. See Section 1.3.7.

but be dominated in the restricted game in which some dominated actions have already been removed. Finally, a game in which this process leaves only one action left for each player is a dominance-solvable game. The only surviving action profile is the unique (pure-strategy) Nash equilibrium of the game.

Potential games. A potential function maps all action profiles (i, j), $i, j \in 1, \ldots, N$, to real numbers (Monderer and Shapley, 1996b). For a two-player game with N actions per player, the potential function can be represented as a $N \times N$ *potential matrix* V whose entries V_{ij} are generic real numbers. A two-player game is a potential game if all differences in payoffs can be expressed through the potential function, i.e. $\forall i, j, k \in 1, \ldots, N$:

$$\begin{aligned} \Pi^R_{ij} - \Pi^R_{kj} &= V_{ij} - V_{kj}, \\ \Pi^C_{ji} - \Pi^C_{ki} &= V_{ij} - V_{ik}. \end{aligned} \tag{1.8}$$

An important example of potential games is congestion games (Rosenthal, 1973; Smith, 1984), which can be used as stylized models of traffic (on roads or on the internet). Potential games certainly have at least one pure Nash equilibrium, and can have multiple (pure and mixed) ones.

Supermodular games. Assume that it is possible to define an order over actions i, k and j, l of players Row and Column respectively. For example, the order could arise by interpreting actions as level of effort. Without loss of generality, consider pairs of actions $i < k$ and $j < l$. A supermodular game (Topkis, 1979; Vives, 1990) is such that

$$\Pi^\mu_{kl} - \Pi^\mu_{il} \geq \Pi^\mu_{kj} - \Pi^\mu_{ij}, \forall \mu \in 1, 2, \forall i < k, j < l \in 1, \ldots, N. \tag{1.9}$$

The intuition for the condition above is that one player's marginal payoff increases if the other player "increases" her action (recall that actions are ordered). This is the concept of *strategic complementarities*, originally formulated for Cournot oligopoly games (Bulow et al., 1985; Cooper and John, 1988).[9] Supermodular games can have multiple (pure and mixed) Nash equilibria, and they certainly have at least one pure Nash equilibrium.

Coordination games. If the actions of a game can be indexed in such a way that all action profiles in which both players play actions with the same index are strict (pure-strategy) Nash equilibria, such a game is a coordination game (Foster and Young, 1998):

$$\Pi^R_{ii} > \Pi^R_{ki}, \Pi^C_{jj} > \Pi^C_{jk}, \forall i, j, k \in 1, \ldots, N. \tag{1.10}$$

[9] This concept is also the starting point for the endogenous business cycle model that we will analyze in the second part of this thesis.

In other words, if all action profiles on the main diagonal are (pure-strategy) Nash equilibria then the game is a coordination game. A typical coordination game is Battle of the Sexes, in which the two members of a couple have to unilaterally decide between attending the opera and going to a football match: The payoffs for both members to coordinate on the same activity are higher than the payoffs for doing different activities. By definition coordination games have multiple pure strategy Nash equilibria. They also typically have multiple mixed ones.

1.3 Convergence properties of some learning algorithms

We define and review the convergence properties of five learning algorithms (reinforcement learning, fictitious play, replicator dynamics, experience-weighted attraction, level-k learning) that have support from experiments with human players, also discussing the experimental evidence. The discussion of these algorithms is preceded by the description of another, simpler, learning rule, best reply dynamics (see Section 1.3.1). Indeed, because a key result of this thesis (see Chapter 3) is that the convergence properties of best reply dynamics are highly correlated to those of the five learning algorithms above, we start upfront with best reply dynamics to develop intuition about the properties of games that may favor (non)convergence. Finally, in Section 1.3.7 we briefly summarize other learning algorithms that have been considered in the literature but are not explicitly studied in this thesis.

1.3.1 Best reply dynamics

Description. A *best reply* is the action that gives the best payoff in response to a given action by an opponent. Under *best reply dynamics* each player myopically responds with the best reply to the opponent's last action. We consider a particular version of best reply dynamics in which the two players alternate moves, each choosing her best response to her opponent's last action.

To see the basic idea consider the game with $N = 4$ shown in Figure 1.1A. Suppose we choose $(1, 1)$ as the initial condition. Assume Column moves first, choosing action $S^C = 2$, which is the best response to Row's action $S^R = 1$. Then Row's best response is $S^R = 2$, then Column moves $S^C = 1$, etc. This traps the players in the cycle $(1, 1) \to (1, 2) \to (2, 2) \to (2, 1) \to (1, 1)$, corresponding to the red arrows. We call this a *best reply 2-cycle*, because each player moves twice. This cycle is an attractor, as can be seen by the fact that starting at $(3, 2)$ with a play by Row leads to the cycle. The first mover can be taken randomly; if the players are on a cycle, this makes no difference, since at most one player has incentive to deviate from the current situation. But when off an attractor the order of the moves can be important. In fact for this example there are two attractors: If we begin

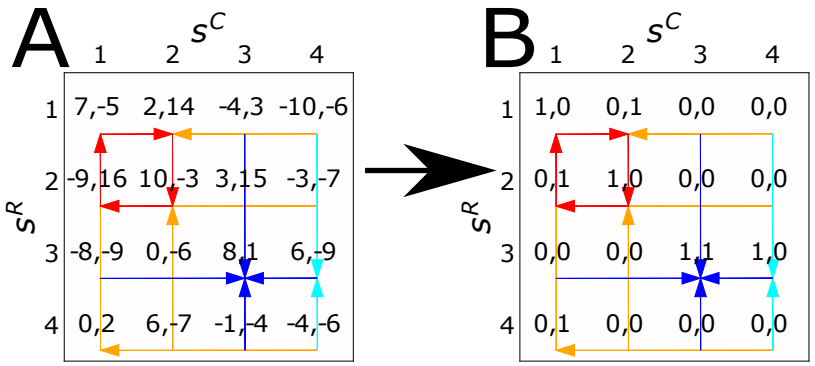


Figure 1.1: **Illustration of best reply dynamics.** $S^R = \{1, 2, 3, 4\}$ and $S^C = \{1, 2, 3, 4\}$ are the possible actions of players Row and Column and each cell in the matrix represents their payoffs (Row is given first). The best response arrows point to the cell corresponding to the best reply. The vertical arrows correspond to player Row and the horizontal arrows to player Column. The arrows are colored red if they are part of a cycle, orange if they are not part of a cycle but lead to one, blue if they lead directly to a fixed point, and cyan if they lead to a fixed point in more than one step. The best reply matrix in B is a Boolean reduction that is constructed to have the same best reply structure as the payoff matrix in panel A, but to only have one and zero as its entries.

at $(3, 2)$ with a play by Column we will arrive in one step at the *best reply fixed point* at $(3, 3)$ (shown in blue). A fixed point of the best reply dynamics is a pure strategy Nash equilibrium.

We characterize the set of attractors of best reply dynamics in a given $N \times N$ payoff matrix Π by a *best reply vector* $\boldsymbol{v}(\Pi) = (n_N, \ldots, n_2, n_1)$, where n_1 is the number of fixed points, n_2 the number of 2-cycles, etc. For instance $\boldsymbol{v} = (0, 0, 1, 1)$ for the example in Figure 1.1.

It will be useful in the following to reduce the payoff matrix to a *best reply (bi-)matrix* as shown in Figure 1.1B. This is done by replacing all best replies for each player by one and all other entries by zero. The best reply matrix has the same best reply structure as the payoff matrix it is derived from, but it ignores any other aspect of the payoffs.

We will show in Chapter 3 that the convergence properties of best reply dynamics are strongly correlated to those of the five learning algorithms described in this chapter. This is because the arrangement of the best replies in the payoff matrix—that we call *best reply structure*—forms a first-order skeleton, characterizing the convergence properties of the game the players are trying to learn. We will see that the presence of best reply cycles is highly correlated to non-convergence, while the presence of best reply fixed points has the opposite effect. In Figure 1.2 we show instances of four of the five classes of games defined in Section 1.2.3 and commonly studied in the literature. While the specific games shown in Figure 1.2 do not contain best reply cycles, it follows directly from definitions (1.7)–(1.10) that these classes of games cannot contain cycles. These results are anticipated here to guide the reader in the literature review in the following sections. As it will be clear, most

proofs for convergence only apply to acyclic games, while examples of non-convergence are most often in cyclic games with no best reply fixed points.

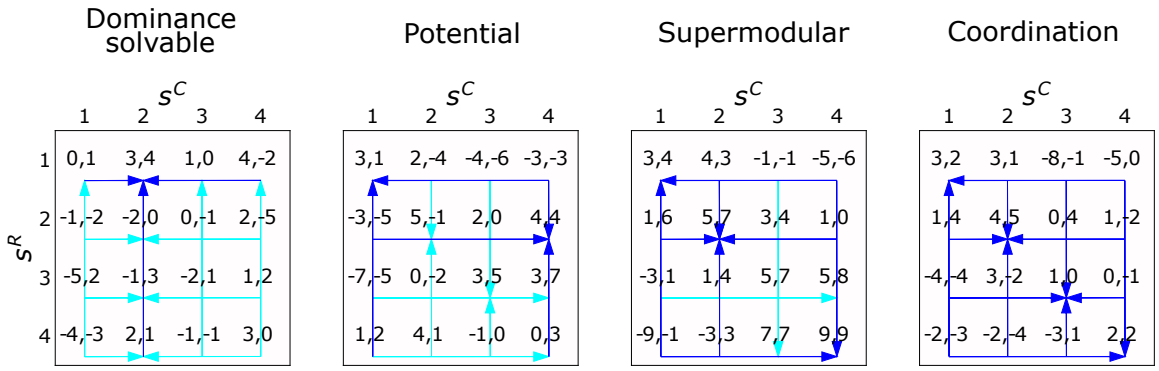


Figure 1.2: **Best reply structure of instances of some extensively studied classes of games**, defined in Section 1.2.3. All these classes of games are best reply acyclic, meaning that they cannot contain best reply cycles.

Background. Best reply dynamics dates back to Cournot (1838), who introduced it as a simple adjustment rule in a model of oligopolistic competition. Because it can be expected that real players use more sophisticated forms of reasoning and it is trivial to characterize its convergence properties in a given game, there are few papers that only focus on best reply dynamics. It is instead common to introduce it to develop intuition about learning and then to focus on other more sophisticated learning algorithms in the rest of the paper. Some papers mostly based on best reply dynamics are Goemans et al. (2005), who use the strongly connected components of a best reply graph as a definition of "sink equilibria", and Hopkins (1999) and Hofbauer and Sorin (2006), who study the continuous-time version.

1.3.2 Reinforcement learning

Reinforcement learning is based on the idea that players are more likely to play actions that yielded a better payoff in the past. It is the standard learning algorithm that is used with limited information and/or without sophisticated reasoning, such as in animal learning. Unfortunately, there exist a plethora of different mathematical specifications of reinforcement learning and results that are supposedly valid for "reinforcement learning" in fact only apply to a given specification. Here we consider two specifications of reinforcement learning. The first is the Bush-Mosteller learning algorithm (Bush and Mosteller, 1955), using the specification of Macy and Flache (2002) and the deterministic approximation introduced in Galla and Farmer (2013). The second specification is introduced by Erev and Roth (1998).

Bush-Mosteller reinforcement learning. In the Bush-Mosteller version of reinforcement learning, each player has a certain level of *aspiration* that corresponds to his discounted average payoff. This leads to a *satisfaction* for each action – positive if the payoff the player gets as a consequence of choosing that action is larger than the aspiration level, negative otherwise. The probability to repeat a certain action is increased if the satisfaction was positive, decreased if it was negative.

More formally, let $A^\mu(t)$ be the aspiration level for player μ at time t. It evolves according to

$$A^\mu(t+1) = (1-\alpha)A^\mu(t) + \alpha\Pi^\mu(s^\mu(t), s^{-\mu}(t)). \tag{1.11}$$

Aspiration is a weighted average of the payoff received at time t, $\Pi^\mu(s^\mu(t), s^{-\mu}(t))$, and past aspiration levels. Therefore, payoffs received in the past are discounted by a factor $1-\alpha$. Here α represents the rate of *memory loss*. Satisfaction is defined by

$$\sigma_i^\mu(t) = \frac{\Pi^\mu\left(i, s^{-\mu}(t)\right) - A^\mu(t)}{\max_{i,j}\left|\Pi_{ij}^\mu - A^\mu(t)\right|}. \tag{1.12}$$

After taking action i at time t, player μ has a positive satisfaction if the payoff he received is higher than his aspiration. Note that α is also called *habituation*, as a repeated choice of action i by player μ leads the aspiration level to correspond to the payoff for action i. Satisfaction would then approach zero, as the player gets habituated. In Eq. (1.12), the denominator is a normalization factor that keeps σ within -1 and +1 (Macy and Flache, 2002). The probability to play action i again is updated as

$$x_i^\mu(t+1) = \begin{cases} x_i^\mu(t) + \beta\sigma_i^\mu(t)(1-x_i^\mu(t)), & \sigma_i^\mu(t) > 0, \\ x_i^\mu(t) + \beta\sigma_i^\mu(t)x_i^\mu(t), & \sigma_i^\mu(t) < 0. \end{cases} \tag{1.13}$$

In the above equation β is the learning rate. Positive satisfaction leads to an increase of the probability (but habituation slows and eventually stops the rise, because habituation decreases the satisfaction), negative satisfaction has the opposite effect. The probabilities for the actions that were not taken are updated from the normalization condition. Denoting them by $j \neq i$, we have

$$x_j^\mu(t+1) = \begin{cases} x_j^\mu(t) - \beta\sigma_i^\mu(t)x_j^\mu(t), & \sigma_i^\mu(t) > 0, \\ x_j^\mu(t) - \beta\sigma_i^\mu(t)\frac{x_i^\mu(t)x_j^\mu(t)}{1-x_i^\mu(t)}, & \sigma_i^\mu(t) < 0. \end{cases} \tag{1.14}$$

The learning algorithm described so far is stochastic. We wish to take a deterministic limit in which the players observe a large sample of actions by their opponent before updating their mixed strategies. We assume that the sample is large enough so that it can be identified with the mixed strategy vector. (See Section 1.3.5 for a discussion on this assumption.) For simplicity, we switch to the notation in which $x_i^R(t) \equiv x_i(t)$ and $x_j^C(t) \equiv y_j(t)$.

We also only consider player Row, because the learning algorithm for Column is equivalent. In the deterministic approximation aspiration updates as

$$A^R(t+1) = (1-\alpha)A^R(t) + \alpha \sum_{i,j} x_i(t)\Pi^R_{ij} y_j(t). \tag{1.15}$$

Satisfaction is calculated for all actions i which are played with positive probability:

$$\sigma^R_i(t) = \frac{\sum_{ij} x_i(t)y_j(t)\left(\Pi^R_{ij} - A^R(t)\right)}{\max_{i,j}|\Pi^R_{ij} - A^R(t)|}. \tag{1.16}$$

Finally, all components of the mixed strategy vector are updated both as if they were played, or as if they were not played, depending on the probabilities $x_i(t)$. The update rule is

$$x_i(t+1) = x_i(t) + x_i(t)\Delta x_i(t) + \sum_{j\neq i} x_j(t)\Delta x_{ij}(t). \tag{1.17}$$

Here, $\Delta x_i(t)$ is the contribution due to the choice of action i by player Row (which occurs with probability $x_i(t)$, hence the multiplying term), and $\Delta x_{ij}(t)$ is the contribution on action i due to the choice of another action j (i.e. the normalization update), each occurring with probability $x_j(t)$. Following Eqs. (1.13) and (1.14), we have

$$\Delta x_i(t) = \begin{cases} \beta\sigma^R_i(t)(1-x_i(t)), & \sigma^R_i(t) > 0, \\ \beta\sigma^R_i(t)x_i(t), & \sigma^R_i(t) < 0, \end{cases} \tag{1.18}$$

and

$$\Delta x_{ij}(t) = \begin{cases} -\beta\sigma^R_j(t)x_i(t), & \sigma^R_j(t) > 0, \\ -\beta\sigma^R_j(t)\frac{x_j(t)x_i(t)}{1-x_j(t)}, & \sigma^R_j(t) < 0. \end{cases} \tag{1.19}$$

We will use this specification of reinforcement learning in Chapter 3.

Other specifications. The Bush-Mosteller model presented above has a number of variants. For example, a simplification used in Cross (1973) and Börgers and Sarin (1997) reads

$$\begin{aligned} x^\mu_i(t+1) &= \Pi^\mu_{ij}(t) + (1-\Pi^\mu_{ij}(t))x^\mu_i(t), \text{ if } i = s^\mu(t), \\ x^\mu_k(t+1) &= (1-\Pi^\mu_{kj}(t))x^\mu_k(t), \forall k \neq s^\mu(t), \end{aligned} \tag{1.20}$$

recalling that $s^\mu(t)$ is the action actually played by μ at t. The disadvantage of this and other similar specifications is that it places strong restrictions on the payoffs. Indeed, for Eq. (1.20) to be valid, $0 < \Pi^\mu_{ij} < 1$ must hold for all μ, i, j. For other specifications it suffices that payoffs be positive (Hopkins, 2002).

All specifications of reinforcement learning listed so far assume that payoffs directly affect mixed strategies. Other specifications assume instead that players update attractions

or propensities only for the actions that are actually played, and then these attractions or propensities are mapped to mixed strategies. For example, letting Q_i^μ denote the attraction to action i for player μ, Erev and Roth (1998) assume that

$$\begin{aligned} Q_i^\mu(t+1) &= Q_i^\mu(t) + \Pi^\mu_{i,s^{-\mu}(t)}, \text{ if } i = s^\mu(t), \\ Q_k^\mu(t+1) &= Q_k^\mu(t), \ \forall k \neq s^\mu(t). \end{aligned} \tag{1.21}$$

They then map the attractions Q_i^μ to probabilities (components of the mixed strategy vector) using a linear mapping: $x_i^\mu(t) = Q_i^\mu(t) / \sum_{k=1}^N Q_k^\mu(t)$. This linear mapping has the main disadvantage that all payoffs must be positive, otherwise the attractions could become negative and the quantities $x_i^\mu(t)$ may not be interpreted as probabilities. This problem is solved in the specifications of Mookherjee and Sopher (1994, 1997) who assume a logit mapping in place of a linear mapping, that is

$$x_i^\mu(t) = \frac{e^{Q_i^\mu(t)}}{\sum_{k=1}^N e^{Q_k^\mu(t)}}. \tag{1.22}$$

This way, even if the propensities become negative, $x_i^\mu(t)$ is always well defined. We will use the specification of reinforcement learning given by Eqs. (1.21)–(1.22) in Chapter 2. As will be discussed in Section 1.3.5, this specification is a special case of Experience-Weighted Attraction.

Background and experimental evidence. Rather than being analyzed in terms of its properties about convergence to equilibrium, reinforcement learning is often studied in terms of its empirical performance on explaining the behavior of real players in experiments. The experiments generally show that reinforcement learning beats Nash equilibrium as a predictive model of human play (Erev and Roth, 1998; Erev et al., 2002, 2007). Moreover, Erev and Roth (1998) and Mookherjee and Sopher (1994, 1997) show that time-averaged play is usually close to Nash equilibrium in games with two actions per player ($N = 2$), but is farther away in games with $N = 4$ and $N = 6$. Note that all these experiments are run with constant-sum games or with other "cyclic" games that only have best reply cycles and no best reply fixed point (see Section 1.3.1). These studies generally bring support to reinforcement learning as compared to fictitious play, described in the next section.

1.3.3 Fictitious play

Fictitious play was first proposed as an algorithm to calculate the Nash equilibria of a game, and later interpreted as a learning algorithm (Brown, 1951; Robinson, 1951). It is an example of a *belief learning* algorithm. In this class of learning algorithms, instead of

learning based on the experienced payoffs as in reinforcement learning, the players update their beliefs on what action could be taken by their opponent, and react to their beliefs.

In fictitious play, each player takes the empirical distribution of actions by her opponent as an estimate of his mixed strategy, calculates the expected payoff of her actions given this belief, and chooses the action that maximizes her expected payoff. In the standard fictitious play algorithm the players weigh all past actions equally and choose the best performing action with certainty. Variants include *weighted fictitious play*, in which the players discount past actions of their opponent and give higher weight to more recent actions, and *stochastic fictitious play*, in which the players select the best performing action with a relatively high probability, and potentially all other actions with a smaller probability (Fudenberg and Levine, 1998). A combination of the two variants leads to *weighted stochastic fictitious play*.

Description. Player Row calculates the j-th component of the expected mixed strategy of Column at time T, which we denote by $\tilde{y}_j(T)$, simply as the fraction of times that j has been played in the past:

$$\tilde{y}_j(T) = \frac{\sum_{t=1}^{T} I(j, s^C(t))}{T}. \tag{1.23}$$

In the above equation, $I(a, b)$ is the indicator function, $I(a, b) = 1$ if $a = b$ and $I(a, b) = 0$ if $a \neq b$. Player Row then selects the action that maximizes the expected payoff at time T,[10]

$$i(T) = \underset{k}{\operatorname{argmax}} \sum_j \Pi^R_{kj} \tilde{y}_j(T). \tag{1.24}$$

The behavior of Column is equivalent. In the weighted fictitious play variant, the update of beliefs is modified to give higher weight to more recent observations:

$$\tilde{y}_j(T) = \frac{\sum_{t=1}^{T} (1-\alpha)^{T-t} I(j, s^C(t))}{\sum_{j=1}^{N} \sum_{t=1}^{T} (1-\alpha)^{T-t} I(j, s^C(t))} = \alpha \frac{\sum_{t=1}^{T} (1-\alpha)^{T-t} I(j, s^C(t))}{(1-(1-\alpha)^T)\, T}, \tag{1.25}$$

where α is the rate of *memory loss*. The action choice follows instead Eq. (1.24), as in standard fictitious play. In the stochastic fictitious play variant, it is instead the action choice that is modified. (The update of beliefs follows Eq. 1.23.) Instead of choosing the action that maximizes the expected payoff, the players select any action with a probability that depends on the expected payoff. The most common mapping from expected payoffs to

[10] In this thesis we study games with generic payoffs, so for typical values at most one action can maximize the expected payoff in Eq. (1.24). Therefore we do not have the problem of choosing one among two or more actions in case of multiple maxima in Eq. (1.24). If that was the case, a common solution is to choose one action at random with uniform probability (Fudenberg and Levine, 1998).

probabilities is through a logit function, as in Eq. (1.22) above. Under this mapping, the component i of the mixed strategy of Row is

$$x_i(T) = \frac{e^{\beta \sum_j \Pi^R_{ij} \tilde{y}_j(T)}}{\sum_k e^{\beta \sum_j \Pi^R_{kj} \tilde{y}_j(T)}}, \tag{1.26}$$

where β is often called *intensity of choice* or *payoff sensitivity*.[11] Following this determination of $x_i(T)$, some studies assume that players actually take an action randomly sampled from their mixed strategy vector, while other studies take a deterministic approximation. The combination of Eqs. (1.25) and (1.26) leads to weighted stochastic fictitious play.

Theoretical results. Fictitious play is probably the learning algorithm whose convergence properties have been investigated most extensively. Robinson (1951) showed that it converges to Nash equilibria in zero-sum games. Since then, convergence of fictitious play has been proved in several other classes of games. Miyazawa (1961) proved that fictitious play converges to (one of) the NE in all 2×2 games[12]—games with two players and two actions available per player. Milgrom and Roberts (1991), Krishna (1992), Hahn (1999) and others proved convergence in some subclasses of supermodular games (there exists no proof that supermodularity alone is sufficient), and Monderer and Shapley (1996a) proved convergence in potential games. Candogan et al. (2013) show that fictitious play also converges in games that are "close" to potential games. Some papers find instead games in which fictitious play does *not* converge. The most famous one is Shapley (1964), who studies a cyclic game with three action per player in which fictitious play cycles with an exponentially increasing period.[13] Another example is Foster and Young (1998): Starting from the premise that fictitious play is believed to converge in most coordination games, the authors construct a few counterexamples in which it fails to converge.

All the research in the previous paragraph was about standard fictitious play. Its variants have been extensively investigated too. Weighted fictitious play can converge (locally) to pure strategy Nash equilibria (Fudenberg and Levine, 1998), yet it fails to converge to mixed strategy ones (Stahl, 1988). Stochastic fictitious play does not converge to Nash equilibria. Instead, it converges to Quantal Response Equilibria (Fudenberg and Kreps, 1993; Benaïm and Hirsch, 1999; Hofbauer and Sandholm, 2002), a generalization of Nash equilibria in which the players choose non-equilibrium actions with a smaller probability (McKelvey and Palfrey, 1995). Cheung and Friedman (1997) are the first to analyze weighted stochastic

[11] In the limit $\beta \to \infty$ the players choose with certainty the action with the largest expected payoff, so we recover standard fictitious play. In the limit $\beta \to 0$ the players choose randomly, disregarding the expected payoff.

[12] Monderer and Sela (1996) showed that the games need to be non-degenerate.

[13] Sparrow et al. (2008) show that a similar game can generate chaotic fictitious play.

fictitious play. Although their paper is mostly focused on the experimental evidence, they show a bifurcation between stability and instability of fixed points in the interior of the strategy space. Benaïm et al. (2009) study weighted stochastic fictitious play too, but take the limits of infinite memory and infinite intensity of choice($\alpha \to 0$, $\beta \to \infty$), and find convergence to a solution concept they propose.

Experimental evidence. The classical experiments that test fictitious play as a learning algorithm that real players use are in Cheung and Friedman (1997). The authors estimate the α and β parameters for each player in a set of coordination, dominance-solvable and cyclic 2×2 games. The paper finds considerable heterogeneity across parameter estimates, with some players best described by short memory ($\alpha \approx 1$), some players by long memory ($\alpha \approx 0$) and the other players in between. The experiment thus does not clarify whether weighted fictitious play improves over standard fictitious play in describing the behavior of real players in these games. Another experiment that tests fictitious play is Van Huyck et al. (1997), but it tests fictitious play only indirectly. The authors show that real players converge to one or the other of the two pure strategy Nash equilibria in a supermodular game, in line with simulations of fictitious play.

1.3.4 Replicator dynamics

Replicator dynamics (Maynard Smith, 1982) is the standard tool used in evolutionary game theory (Hofbauer and Sigmund, 1998). It is a stylized model representing the evolution of individuals with certain traits in a population. In this model the fitness of each trait depends on the population shares of the other traits, and on the average fitness. Although replicator dynamics is mostly used in population biology, it has also been studied as a learning algorithm in game theory. The key connection is through the *population of ideas* (Börgers and Sarin, 1997). According to this interpretation, each action can be viewed as a trait, and the evolution of the population shares of the traits corresponds to the dynamics of the components of the mixed strategy vector.

The most typical form of replicator dynamics concerns only one population. This only applies to symmetric games, which can be seen as between a focal player and the rest of the population. However, since in this thesis we are interested in generic games, asymmetric payoff matrices will typically occur. This naturally leads to *two-population* replicator dynamics. The dynamical properties of the two-population version are different from those of the one-population algorithm. For our purposes, the most important difference is that one-population replicator dynamics can converge to mixed strategy Nash equilibria, whereas

two-population replicator dynamics only converges to *strict* Nash equilibria (i.e. pure strategy Nash equilibria in which the payoff at equilibrium is strictly larger than any other payoff that can be obtained if the opponent does not change his action) (Gintis, 2000).

Description. Letting x_i and y_j denote the population shares of individuals with traits i and j respectively, two-population replicator dynamics reads

$$\begin{aligned} \dot{x_i}(t) &= x_i(t) \left(\sum_j \Pi^R_{ij} y_j(t) - \sum_{kj} x_k(t) \Pi^R_{kj} y_j(t) \right), \\ \dot{y_j}(t) &= y_j(t) \left(\sum_i \Pi^C_{ji} x_i(t) - \sum_{ik} y_k(t) \Pi^C_{ki} x_i(t) \right). \end{aligned} \tag{1.27}$$

The shares of trait i in population Row and trait j in population Column evolve according to the fitness of that trait (as given by the expected payoff) compared to the average fitness in the respective population (Hofbauer and Sigmund, 1998).

We will perform simulations of replicator dynamics, for which we will use the discrete-time version proposed by Maynard Smith (Maynard Smith, 1982, Appendix D, p. 183),

$$\begin{aligned} x_i(t+1) &= x_i(t) \frac{1 + \delta \sum_j \Pi^R_{ij} y_j(t)}{1 + \delta \sum_{kj} x_k(t) \Pi^R_{kj} y_j(t)}, \\ y_j(t+1) &= y_j(t) \frac{1 + \delta \sum_i \Pi^C_{ji} x_i(t)}{1 + \delta \sum_{ik} y_k(t) \Pi^C_{ki} x_i(t)}. \end{aligned} \tag{1.28}$$

where δ is a parameter that should take small values. (In his notation, Maynard Smith (Maynard Smith, 1982, Appendix D, p. 183) uses a parameter $C = 1/\delta$.)

Theoretical results. The local convergence properties of replicator dynamics are well-understood. Only Evolutionary Stable Strategies (ESS) are locally stable (Maynard Smith, 1982), so when it converges to a fixed point, replicator dynamics converges to one of them. ESS for one-population replicator dynamics can be mixed or pure strategy Nash equilibria; for the two-population version, ESS can only be pure Nash equilibria. Global convergence has only been proved in dominance-solvable games (Nachbar, 1990; Hofbauer and Weibull, 1996). Convergence to periodic cycles is well-known to occur in cyclic games such as Matching Pennies and Rock Paper Scissors (Maynard Smith, 1982); in other similar games convergence to chaotic attractors has also been shown to be possible (Skyrms, 1992; Chawanya, 1995; Sato et al., 2002). A recently proposed general solution concept for replicator dynamics, which encompasses both Nash equilibria and unstable dynamics, is chain recurrent sets (Papadimitriou and Piliouras, 2016).

Another set of theoretical results that will be useful in this thesis (see Appendix 2.A.1) concerns the equivalence between replicator dynamics and some specification of reinforcement learning (Börgers and Sarin, 1997; Hopkins, 2002). Moreover, replicator dynamics can be seen as a special case of Experience-Weighted Attraction (see Section 1.3.5). Obtaining replicator dynamics from Experience-Weighted Attraction leads to a finite memory generalization of the replicator equations (Sato and Crutchfield, 2003), with different convergence properties.

Experimental evidence. Replicator dynamics has mostly been used as a stylized model of evolution in population biology and ecology, yet some experiments with human players have tested it as a behavioral model. Van Huyck et al. (1995) consider two coordination games and show that players end up in the same equilibria as those predicted by replicator dynamics. Remarkably, the size of the basins of attraction of replicator dynamics has predictive power over which equilibrium is actually selected. Friedman (1996) comes to similar conclusions in a wider set of coordination and dominance-solvable games. He also considers a cyclic game—Buyer-Seller—reporting different outcomes in different experimental sessions. In some cases play spirals outward from the equilibrium, in others it spirals towards the equilibrium, in others it circles around the equilibrium neither converging nor diverging, and in one case there is no clear pattern. Cason et al. (2013) design an experiment specifically to test for cycles in another cyclic game, Rock-Paper-Scissors. They find evidence that players indeed cycle, as predicted by replicator dynamics.

1.3.5 Experience-Weighted Attraction

Experience-Weighted Attraction (EWA) has been proposed by Camerer and Ho (1999) to generalize reinforcement and belief learning algorithms (such as fictitious play, or best reply dynamics). The connection between reinforcement and belief learning lies in the update of the actions that are not played, i.e. in considering *forgone payoffs*. If only the probabilities of the actions that are played are updated, EWA reduces to the specification of reinforcement learning given by Eqs. (1.21)–(1.22). If all probabilities are updated with the same weight, EWA reduces to fictitious play or best reply dynamics, depending on the memory parameter. Finally, note that EWA also reduces to replicator dynamics by taking the limits of some parameters (e.g., by taking the limit of infinite memory), as shown e.g. in Sato and Crutchfield (2003).

Description. In EWA, the mixed strategies are determined from the so-called *attractions* or *propensities* $Q_i^\mu(t)$. These are real numbers that quantify the level of appreciation of

player μ for action i at time t. Attractions are not normalized, so the mapping to probabilities occurs through a logit,

$$x_i^\mu(t+1) = \frac{e^{\beta Q_i^\mu(t+1)}}{\sum_j e^{\beta Q_j^\mu(t+1)}}, \tag{1.29}$$

where β has the same interpretation as in Eq. (1.26). A similar expression holds for $y_j(t+1)$. The propensities update as follows:

$$Q_i^\mu(t+1) = \frac{(1-\alpha)\mathcal{N}(t)Q_i^\mu(t) + (\delta + (1-\delta)I(i, s^\mu(t+1))\Pi^\mu(i, s^{-\mu}(t+1))}{\mathcal{N}(t+1)}, \tag{1.30}$$

where

$$\mathcal{N}(t+1) = (1-\alpha)(1-\kappa)\mathcal{N}(t) + 1. \tag{1.31}$$

Here $\mathcal{N}(t)$ represents *experience* because it increases monotonically with the number of rounds played; the more it grows, the smaller becomes the influence of the received payoffs on the attractions (as the denominator of (1.30) increases). The propensities change according to the received payoff when playing action i against action $s^{-\mu}$ by the other players, i.e. $\Pi^\mu(i, s^{-\mu}(t+1))$. The indicator function $I(i, s^\mu(t+1))$ is equal to 1 if i is the actual action that was played by μ at time $t+1$, that is $i = s^\mu(t+1)$, and equal to 0 otherwise. All attractions (those corresponding to actions that were and were not played) are updated with weight δ, while an additional weight $1-\delta$ is given to the specific attraction corresponding to the action that was actually played. Finally, the memory loss parameter α determines how quickly previous attraction and experience are discounted and the parameter κ interpolates between cumulative and average reinforcement learning (Camerer and Ho, 1999).

The update of the attractions (1.30) is stochastic, because it depends on which actions $s^\mu(t)$ and $s^{-\mu}(t)$ are randomly sampled from the mixed strategy vectors of the two players at time t. As with Bush-Mosteller reinforcement learning (see Section 1.3.2), we take a deterministic limit (see below for a discussion on this assumption). In the limit, Eq. (1.30) reads

$$Q_i^R(t+1) = \frac{(1-\alpha)\mathcal{N}(t)Q_i^R(t) + (\delta + (1-\delta)x_i(t))\sum_j \Pi_{ij}^R y_j(t)}{\mathcal{N}(t+1)}, \tag{1.32}$$

and a similar expression holds for Column.

The deterministic limit corresponds to the assumption that the players play against each other an infinite number of times before updating their attractions. The empirical frequency of their actions corresponds then to their mixed strategy. This sort of argument was already made by Crawford (1974) and justified by Conlisk (1993a) in terms of "two-rooms experiments": The players only interact through a computer console and need to

specify several actions before they know the actions of their opponent.[14] Another analogy that helps interpret this assumption is the difference between *batch* and *online* in machine learning. With batch learning, the parameters of the machine learning model are updated after computing the performance score of the model on the whole dataset. With online learning, the parameters are instead updated for any observed data point. In this analogy, batch learning corresponds to the deterministic approximation (1.32) and online learning to the stochastic version (1.30).

The deterministic assumption is useful from a theoretical point of view and does not affect the results in most cases; the only difference when noise is allowed is a blurring of the dynamical properties. To show this point explicitly, in the analysis in Chapters 2 and 3 we will consider both the deterministic and stochastic versions of EWA.

Theoretical results. EWA has mostly been used to describe the behavior of real players in experiments, with little attention to its theoretical properties. To the best of our knowledge, the first (approximate) analysis of its convergence properties, in two-player games with many actions available per player, is in Galla and Farmer (2013). Sato and Crutchfield (2003) derive a continuous-time limit of a learning algorithm that essentially coincides with EWA. They show that the continuous-time version of their algorithm is a generalized replicator dynamics with finite memory. In Appendix 2.A.1 we provide an alternative derivation.

Experimental evidence. Camerer and Ho (1999) compare the out-of-sample predictive power of EWA with that of pure reinforcement learning or pure belief learning (i.e. weighted stochastic fictitious play), reporting the best performance for EWA in most cases. According to the authors, the reason is that real players use information about experienced payoffs, as in reinforcement learning, but they also try and predict the next actions of their opponent, as in belief learning. Since EWA interpolates between these two extremes, it is a better description of the behavior of real players.

1.3.6 Level-k learning

All learning algorithms described so far are backward looking, in the sense that players react to the last strategy of their opponent, with the implicit assumption that this will also be played on the following iteration of the repeated game. In some situations this might be implausible, and we may wish to consider forward-looking players. We refer to level-k

[14] Bloomfield (1994) implemented this idea in an experimental setup. Cheung and Friedman (1997) also consider a matching protocol and a population setting in which the players are matched with all players from the other population. This has a similar effect in eliciting mixed strategies.

learning as a generalization of anticipatory learning, as proposed by Selten (1991). Selten assumed that player Row does not believe that Column would behave as she did in the past. He rather tries to outsmart her by best replying to the strategy that he thinks she will play on the following time step. Row's model of the future strategy of Column is given by her learning algorithm, that is assumed to be known by Row: he best replies to the strategy that Column would play under her learning algorithm.

This idea can be generalized by assuming that the players can think k steps ahead (Lecutier, 2013; Evans, 2013). In level-k thinking (Nagel, 1995; Crawford et al., 2013) k-players assume that the other players are level $k-1$, and the process is iterated down to level 1. Level-1 players choose randomly. Level-2 players know that level-1 players choose randomly, and select the strategy that yields the highest payoff given this piece of information. Level-3 players know how level-2 players behave, and react accordingly, and so on.

In our case, level-1 players are EWA learners. Level-2 players know that level-1 players update their strategies using EWA, and try to get a better payoff by pre-empting their opponent's strategy. Level-3 players would know how level-2 players choose their strategy, and select the best possible strategy in response. Here we will assume that both players are level-2, as we did not find a substantial difference with larger values of k (which quickly become behaviorally implausible).

Description. For convenience, we combine Eqs. (1.29) and (1.32):

$$x_i^1(t+1) = \frac{x_i(t)^{(1-\alpha)\mathcal{N}(t)/\mathcal{N}(t+1)} \exp\Big(\beta(\delta + (1-\delta)x_i(t)) \sum_j \Pi_{ij}^R y_j(t)/\mathcal{N}(t+1)\Big)}{Z_x(t+1)}, \quad (1.33)$$

with $Z_x(t+1) = \sum_l x_l(t)^{(1-\alpha)\mathcal{N}(t)/\mathcal{N}(t+1)} \exp\Big(\beta(\delta + (1-\delta)x_l(t)) \sum_j \Pi_{lj}^R y_j(t)/\mathcal{N}(t+1)\Big)$. We are using superscript 1 to indicate that player Row is a level-1 (i.e. an EWA) learner. A similar expression holds for Column.

We denote the right-hand side in Eq. (1.33) by $F(y(t))$, with $y(t) = (y_1(t), \dots y_N(t))$. So, $x_i^1(t+1) = F(y(t))$: A level-1 player Row learns based on the past mixed strategy vector of Column. We define

$$y_j^2(t+1) = F(x^1(t+1)). \quad (1.34)$$

Here Column is a level-2 player as she believes that Row is a level-1 player, and therefore updates his strategies using Eq. (1.33). In general,

$$y_j^k(t+1) = F(x^{k-1}(t+1)). \quad (1.35)$$

Theoretical results. In some cases, increasing the level of the players results in more stable learning dynamics. For example, Selten (1991) considered a learning algorithm, proposed in Crawford (1974), that cannot reach mixed strategy Nash equilibria. However, Selten shows, having both players level-2 makes it possible for the algorithm to reach mixed equilibria, for some parameter values. Similar results are obtained in the master theses of Lecutier (2013) and Evans (2013), who study level-k EWA. They find that increasing the level of the players increases the size of the parameter space that leads to convergence. However, increasing the level of reasoning could in principle result in less convergence (Huberman and Hogg, 1988). In generic games, we will show, increasing the level of the players has negligible effect on the likelihood of convergence.

Experimental evidence. There exists abundant experimental evidence on level-k reasoning in games (Crawford et al., 2013). Some experiments involve learning (Nagel, 1995; Stahl, 1996), but this is usually intended in the sense that level-k predictions are possible actions that can be reinforced. For example, in Nagel (1995) players can make level-1, level-2, etc., predictions and learn to choose the predictions that lead to highest payoffs. There is smaller experimental evidence of level-k learning intended in the sense of this section. Tang (2001) provides some experimental evidence in favor of Selten (1991)'s model. Camerer et al. (2002) consider an EWA model in which level-2 players hold beliefs on the fraction of level-2 vs. level-1 players. They show that this "sophisticated" model beats an equivalent model in which all players are level-1 in predicting data from p-beauty contest games (Nagel, 1995). All experimental results listed so far are about how well level-k learning describes behavior of real players. Frey and Goldstone (2013), instead, conduct an experiment on a multiplayer generalization of Rock-Paper-Scissors and show that the combination of iterated reasoning and learning leads to cycles in belief space.

1.3.7 Other learning algorithms

This thesis investigates the convergence properties of the learning algorithms described in Sections 1.3.2 through 1.3.6. Obviously the literature has considered a much wider set of learning algorithms whose convergence properties may be very different from those in the sections above. Our choice of learning algorithms was motivated by the empirical evidence from experiments with human players. Many other algorithms have been proposed that are not meant to describe the behavior of real players, rather to have certain convergence properties *by design*. These algorithms are very useful to find equilibria in applications ranging from mechanism design to engineering and computer science. In this section we

give an account of some of these algorithms, plus other algorithms that do not have as clear an experimental evidence as those in Sections 1.3.2 through 1.3.6.

A class of simple learning algorithms that dates back to Arrow and Hurwicz (1960) is *gradient learning*. Although some details are different, under learning rules in this family the components of the mixed strategy vectors evolve "in the direction" of increasing payoffs. Arrow and Hurwicz (1960) prove that gradient learning converges to pure strategy Nash equilibria, yet Crawford (1974, 1985) and Stahl (1988) show that it does not converge to mixed strategy ones. Selten (1991) and Conlisk (1993a,b) propose amendments to basic gradient learning that make convergence to mixed equilibria possible.

A class of learning rules whose convergence properties are extremely good is *regret learning*. Players using algorithms in this class calculate the missed payoff, or "regret", from not having played other actions in the past, and try to minimize regret. Hart and Mas-Colell (2000) show that these simple adaptive rules converge to correlated equilibria in all games.[15] Foster and Young (2006) propose another regret learning algorithm that instead converges—within a small distance ϵ—to Nash equilibria in all two-player games.[16] However, these algorithms, together with other algorithms proposed by the same authors (Foster and Young, 2003; Young, 2009), make exhaustive searches of the strategy space, and so may take really long to find the Nash equilibrium.

Gradient learning and regret learning are still adaptive processes. A third class of learning rules that is more sophisticated is *Bayesian learning*. Under this class of learning algorithms the players do not know the payoffs of their opponents, but have a prior over what these payoffs could be. At each round of the game they play the Bayesian Nash equilibrium consistent with this prior and update the prior after observing the actions taken by their opponent. Jordan (1991) shows that this process converges to Nash equilibria in all normal form games for myopic players, while Kalai and Lehrer (1993) extend this result to forward-looking players that consider histories of play (i.e. repeated games setting). However, these results assume that players put positive probability on all actions that could be taken by their opponents. Jordan (1993), Nachbar (1997) and Foster and Young (2001) show that this assumption is difficult to satisfy in practice because of a tension between rationality and prediction. Another learning algorithm closely connected to the Bayesian approach is calibrated learning (Foster and Vohra, 1997). This learning rule always converges to correlated equilibria, provided players' predictions are properly "calibrated".

A few other papers are worth mentioning. One is Sonsino (1997), in which two players play best reply dynamics but are able to recognize patterns in their opponent's play. This

[15] See Hart (2005) and Blum and Mansour (2007) for an exhaustive account of the various versions of "regret matching" as first proposed in Hart and Mas-Colell (2000).

[16] Germano and Lugosi (2007) extends this result to an arbitrary number of players

leads the players to pure strategy Nash equilibria. Arieli and Young (2016) consider a better reply dynamics known as "pairwise comparison dynamics". They show that it converges fast to Nash equilibria in all "weakly acyclic" games, that is, games with no best reply cycles. Finally, it is worth mentioning the large literature on Cournot games (Bischi et al., 2009). Although these are formally normal-form games, the continuity of the action set makes the analysis of their convergence properties somewhat different from the approach in this thesis.

1.4 Statistical properties of games

As motivated in the Introduction and in Section 3.2 more thoroughly, in Chapter 3 we study the convergence properties of randomly generated games. While studying the generic properties of randomly generated systems is very common in the natural sciences, it is not common in game theory. However, a few papers exist that characterize the number of Nash equilibria and other properties of randomly generated games. A subset of these papers also studies stability of equilibria. The techniques used range from combinatorics to statistical mechanics.

Number of pure equilibria. Goldberg et al. (1968) consider two-player games in which all payoffs are independently randomly generated. They show that in the limit in which players have an infinite number of actions, $N \to \infty$, the probability that there exists at least one pure strategy Nash equilibrium is $1 - 1/e$, where $e \approx 2.718$ is Euler's number. Dresher (1970) extends this result to multiple players. He finds that the probability of at least one pure equilibrium remains $1 - 1/e$ for any number of players. The analysis of pure equilibria in normal-form games under independent payoffs are completed by Powers (1990) and Stanford (1995, 1996), who study the distribution of the number of Nash equilibria in any ensemble of such games.

It is also interesting to study properties of equilibria when payoffs at strategy profiles are correlated. In two-player games, this means that at any given cell of the payoff matrix the two payoffs to the two players are drawn from a bi-variate distribution with correlation Γ. Goldberg et al. (1968), for example, consider games in which players' interests are purely in conflict, such as zero-sum games ($\Gamma = -1$), or exactly parallel, that is the highest payoffs for both players occur in the same cell of the payoff matrix for any given pair of rows and columns ($\Gamma = 1$). In zero sum games the probability of existence of at least one pure equilibrium tends to zero for $N \to \infty$; in games with perfect correlation at any strategy profile, it tends to one. In this latter class of games, Stanford (1999) shows that the number of pure strategy equilibria diverges to infinity as the number of strategies increases. Rinott

and Scarsini (2000) calculate the expected number of pure equilibria for any value of the correlation $\Gamma \in [-1, 1]$, showing that it increases with Γ.

Number of pure and mixed equilibria. Berg and Weigt (1999) and Berg (2000) study this question using methods from statistical mechanics, while McLennan and Berg (2005) and McLennan (2005) use more rigorous polynomial methods from combinatorics. Berg and Weigt (1999) and Berg (2000) show how the expected number of Nash equilibria varies as a function of the number of available actions to each player N and the correlation of the payoffs Γ. Increasing N and increasing Γ both lead to an exponential increase in the number of equilibria. McLennan and Berg (2005) derive a more precise expression in the case of two-player games, while McLennan (2005) extends this result to games with an arbitrary number of players.

Other properties of the game. Other properties of the games are typically investigated. Berg and Engel (1998) study the value of the game (the expected payoff), and the fraction of actions played with positive probability, at a typical Nash equilibrium, in two-player zero-sum games. They investigate these properties as they vary the number of actions available to player Row, N, with respect to the ones available the player Column, M,[17] and also as they increase the row- or column-correlation of the payoffs. They find that the value of the game decreases and the fraction of actions played with positive probability increases as the ratio M/N increases. Finally, they find that the fraction of actions played with positive probability decreases with the row- or column-correlation of the payoffs. Berg and Weigt (1999) and Berg (2000) study the same two properties (value of the game and fraction of actions played with positive probability) as they vary the correlation of the payoffs at the same site Γ. They find that the value of the game increases and the fraction of actions decreases with increasing Γ.

Cohen (1998) studies a different property, Pareto optimality. In particular, he asks how likely it is that Nash equilibria in randomly generated two-player games are also Pareto optimal, that is, there are no other action profiles that would increase the payoff of one player without decreasing the payoff of another player. He finds that in the limit $N \to \infty$, about 60% of Nash equilibria are also Pareto optimal. On the contrary, with the smallest $N = 2$, about 90% of Nash equilibria are Pareto optimal. This implies, among other things, that a configuration of the payoffs such as the Prisoner Dilemma, a game with $N = 2$ in which the Nash equilibrium is not Pareto optimal, is highly unlikely in randomly generated

[17] The asymmetry between the players originates from the zero-sum assumption. Indeed, Row receives positive payoffs while Column gets negative—and opposite—payoffs.

games. This configuration becomes more likely as $N \to \infty$, yet it only occurs in about 40% of the cases.

Convergence to equilibria. The literature on the convergence properties of learning algorithms in randomly generated games is more scarce. The replicator dynamics has been investigated most extensively, by the evolutionary game theory community. Diederich and Opper (1989) and Opper and Diederich (1992) study the existence and stability of fixed points of replicator dynamics as they vary the self-interaction among species, i.e. the value of the diagonal elements in the payoff matrix. As this quantity increases, the fraction of actions played with positive probability also rises. Lower values of this quantity make the fixed points unstable, making it possible that limit cycles or chaotic solutions arise (Opper and Diederich, 1999). Gokhale and Traulsen (2010) and Han et al. (2012) study related properties in multi-population games, that is evolutionary games with multiple players.

Galla and Farmer (2013) and Sanders et al. (2018) study the convergence properties of Experience-Weighted Attraction (see Section 1.3.5) in two-player (Galla and Farmer, 2013) and multiplayer (Sanders et al., 2018) games, as they vary the number of actions N and the correlation between the payoffs at the same site Γ. They show that convergence becomes less likely with increasing N and decreasing Γ, even more so if the number of players increase. They demonstrate their results with numerical simulations and with statistical mechanics calculations in the $N \to \infty$ limit.

Finally, Amiet et al. (2019) study a version of best reply dynamics in two-action multiplayer games in which players do not play following a fixed sequence (as we assume in Section 1.3.1). They show that this version of best reply dynamics generally converges to a pure equilibrium as the number of players increases and the probability of ties is not too high.

Chapter 2

A taxonomy of learning dynamics in 2x2 games

2.1 Overview

A large literature in behavioral game theory has proposed and experimentally tested various learning algorithms, but a comparative analysis of their equilibrium convergence properties is lacking. Here we provide a "taxonomy" of learning dynamics by analyzing Experience-Weighted Attraction (EWA), a learning algorithm that generalizes fictitious play, best reply dynamics, reinforcement learning and also replicator dynamics. We do so in games with two players and two actions available to each player (2×2). These games are simple enough to allow a full analytical characterization of the asymptotic behavior of EWA learning.

We begin this chapter by some preliminary considerations (Section 2.2). These include introducing some notation and building intuition about the qualitative behavior of learning, before diving into the mathematical analysis. Our challenge is to characterize a 13-dimensional parameter space, composed of the eight payoffs that fully define a 2×2 game, the four learning parameters of EWA, and the choice between deterministic and stochastic dynamics. To address the exploration of the parameter space, we follow a modular strategy.

In Section 2.3, we first focus on a baseline scenario in which only some parameters are considered. In the baseline scenario we characterize the position and stability of fixed points of EWA as a function of combinations of payoffs and learning parameters. We show a bifurcation between convergent and unstable learning dynamics in cyclic 2×2 games in the same family as Matching Pennies. Other results concern the probability of convergence close to Pareto-inefficient pure strategy Nash equilibria and the relation between the convergence properties and well-known classes of 2×2 games.

We then explore the effects of other parameters in Section 2.4. One of the most interesting extensions of the baseline scenario is based on changing a key EWA parameter. This is the weight δ given to payoffs corresponding to actions that were played vs. actions that

were not played (see Section 1.3.5), and interpolates between reinforcement learning and belief learning algorithms such as fictitious play. When players only consider the payoffs corresponding to actions that were played, they converge to "lock-in" fixed points that can be arbitrarily far from Nash equilibria.

Our analysis is useful for various reasons. From an experimental point of view, EWA is used to model behavior in many experiments, so our analysis provides theoretical guidance on whether learning dynamics can be expected to converge to equilibrium. From a theoretical point of view, we generalize results known in the literature for the learning algorithms that are a particular case of EWA, obtaining interesting new insights.

A working paper based on the results presented in this Chapter (Pangallo et al., 2017) is available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2913183`.

2.2 Preliminaries

Before diving into the mathematical analysis, we introduce some additional notation and give an intuitive account of the outcome of learning. In Section 2.2.1, we list several types of 2×2 games. We then specify under which parameter restrictions EWA reduces to the learning algorithms it generalizes (Section 2.2.2). We lay out a plan for the analysis of the parameter space in Section 2.2.3. Finally, in Section 2.2.4 we take advantage of the concepts just introduced to build intuition into the qualitative behavior of learning.

2.2.1 Classes of 2×2 games

Despite being very simple, 2×2 games encompass many of the strategic tensions that are studied by game theorists. In the following, we classify 2×2 games into some classes that correspond to some of these strategic tensions. While these classes are mostly equivalent to the ones listed in Section 1.2.3 for generic normal-form games, for illustrative purposes we find it best to specify them in the context of 2×2 games. We also introduce some combinations of payoffs that will be useful in the following analysis.

The notation in this chapter is the same as the one introduced in Section 1.2.1. For convenience, we only write the generic 2×2 payoff matrix (Eq. 1.1) as

$$\begin{pmatrix} a, e & b, g \\ c, f & d, h \end{pmatrix}, \tag{2.1}$$

where for example $\Pi^R_{11} = a$ and $\Pi^C_{11} = e$. We denote action $i = 1$ of player Row by s^R_1, action $j = 1$ of player Column by s^C_1, and similarly action 2 of the two players by s^R_2 and s^C_2.

Based on the properties of the game one wants to look at, it is possible to construct several classifications of 2×2 games. Perhaps the most famous classification was proposed by Rapoport and Guyer (1966), who constructed all distinct games that can be obtained by ordering the payoffs of the two players in all possible ways. For example, one possible ordering is $a > b > c > d$ and $e > f > g > h$. Excluding equivalent orderings, Rapoport and Guyer (1966) found 78 classes of games that represent different strategic tensions.

Our analysis below shows that we do not need such a fine-grained classification of payoff matrices to build intuition into the outcome of EWA learning dynamics. It is instead enough to consider the pairwise ordering of the payoffs to one player, when the action of the opponent is kept fixed. These comparisons are for example between a and c for player R, when the action of Column is fixed to s_1^C, and b versus d when Column's action is fixed to s_2^C. In principle there are $2^4 = 16$ such pairwise orderings. For our purposes we can group these orderings into 4 classes, as illustrated in Table 2.1.[18]

Class of game	Payoff matrix: $\begin{pmatrix} a,e & b,g \\ c,f & d,h \end{pmatrix}$	Conditions on payoff combinations	Nash Equilibria
Coordination	$a > c$, $b < d$, $e > g$, $f < h$.	$\|B\| < \|A\|$, $\|D\| < \|C\|$, $A > 0$, $C > 0$.	Two pure strategy (s_1^R, s_1^C), (s_2^R, s_2^C) and one mixed strategy NE.
Anticoordination	$a < c$, $b > d$, $e < g$, $f > h$.	$\|B\| < \|A\|$, $\|D\| < \|C\|$, $A < 0$, $C < 0$.	Two pure strategy (s_1^R, s_2^C), (s_2^R, s_1^C) and one mixed strategy NE.
Cyclic	$a > c$, $e < g$, $b < d$, $f > h$; $a < c$, $e > g$, $b > d$, $f < h$.	$\|B\| < \|A\|$, $\|D\| < \|C\|$, $AC < 0$.	Unique mixed strategy NE.
Dominance-solvable	$a > c$, $e > g$, $b > d$, $f > h$, and all other 11 orderings.	$\|B\| > \|A\|$ or $\|D\| > \|C\|$	Unique pure strategy NE.

Table 2.1: Relevant classes of two-person, two-action games and their relation with the payoff combinations that determine the learning dynamics. The classes of games are defined from the pairwise ordering of the payoffs. The payoff combinations $A = (a + d - b - c)/4$, $B = (a + b - c - d)/4$, $C = (e + h - f - g)/4$ and $D = (e + f - g - h)/4$ quantify certain structures in the payoff matrix. In particular, A and C quantify the preference of the players for "diagonal" outcomes and B and D measure the dominance of one of their actions over the other.

Coordination and anticoordination games. These correspond to the orderings $a > c$, $b < d$, $e > g$, $f < h$ (coordination games), and $a < c$, $b > d$, $e < g$, $f > h$ (anticoordination games).[19] Coordination games have two pure strategy NE, one at (s_1^R, s_1^C) and the other at (s_2^R, s_2^C). In addition there is one mixed strategy NE. Two well-known examples of co-

[18] Our classification is relatively standard, for example it is very close to the one in Chapter 10 of Hofbauer and Sigmund (1998).

[19] Anticoordination and coordination games can be seen as equivalent, as each type of game can be obtained from the other by relabeling one action of one player (e.g. rename s_1^R into s_2^R and vice versa). However, part of our analysis will be based on symmetric games, and once the constraint of symmetry is enforced some properties of anticoordination and coordination games become different. Therefore, for clarity of exposition we keep coordination games distinct from anticoordination games.

ordination games are Stag-Hunt and Battle of the Sexes (Osborne and Rubinstein, 1994). Anticoordination games also have two pure strategy and one mixed strategy NE, but at the pure strategy NE the players choose strategies with different labels, i.e. (s_1^R, s_2^C) and (s_2^R, s_1^C). A well-known example of an anticoordination game is Chicken.

Cyclic games. These correspond to the orderings $a > c$, $e < g$, $b < d$, $f > h$ or $a < c$, $e > g$, $b > d$, $f < h$. Games of this type are characterized by a best reply cycle (Section 1.3.1). For example, if one considers the first set of orderings, the best reply to s_1^R is for Column to play s_2^C. In response Row would choose s_2^R, Column would then play s_1^C, and the process would never converge to a fixed point. Cyclic games have a unique mixed strategy NE and no pure strategy NE. The prototypical cyclic game is Matching Pennies, which is a zero-sum game, but cyclic games in general need not be zero- or constant-sum.

Dominance-solvable games. These comprise all 12 remaining orderings. These games have a unique pure strategy NE, which can be obtained via elimination of dominated strategies. For instance, if $a > c$, $e > g$, $b > d$, $f > h$, the NE is (s_1^R, s_1^C). The well-known Prisoners' Dilemma is a 2×2 dominance-solvable game.

For the analysis of the outcome of learning it is useful to introduce the following combinations of payoffs,

$$\begin{aligned} A &= \frac{1}{4}\left(a + d - b - c\right), \\ B &= \frac{1}{4}\left(a + b - c - d\right), \\ C &= \frac{1}{4}\left(e + h - f - g\right), \\ D &= \frac{1}{4}\left(e + f - g - h\right). \end{aligned} \tag{2.2}$$

Broadly speaking, a positive value of the parameter A indicates the preference of player Row for outcomes of the type (s_1^R, s_1^C) or (s_2^R, s_2^C) relative to outcomes (s_1^R, s_2^C) and (s_2^R, s_1^C). Similarly, positive values of C indicate the preference of player Column for the outcomes (s_1^R, s_1^C) or (s_2^R, s_2^C). These action combinations are the ones on the main diagonal of the payoff matrix. If instead A is negative, player Row prefers off-diagonal combinations in the payoff matrix, and similarly player Column prefers off-diagonal combinations when C is negative. The strength of these preferences for diagonal or off-diagonal combinations are determined by the modulus $|A|$ for player Row and by $|C|$ for player Column. The parameter B is a measure for the dominance of player Row's first action over her second, and similarly D measures the dominance of player Column's first action over her second.

The class of a particular game, further to being determined by the pairwise orderings of the payoffs discussed above, can also be established based on A, B, C and D.

Proposition 1. *Consider a 2-player, 2-action game. The following statements hold:*
(i) The game is dominance solvable if $|B| > |A|$ *or* $|D| > |C|$.
(ii) If none of the conditions in (i) hold, and in addition $A > 0, C > 0$ *the payoff matrix describes a coordination game.*
(iii) If none of the conditions in (i) hold, and in addition $A < 0, C < 0$ *the payoff matrix describes an anticoordination game.*
(iv) If none of the conditions in (i) hold, and A *and* C *have opposite signs (i.e.* $AC < 0$*) the game is cyclic.*

Proof of proposition 1. We only prove that we have a coordination game (defined by $a > c$, $e > g$, $d > b$, $h > f$) if and only if $|A| > |B|$, $|C| > |D|$, $A > 0$, $C > 0$. The other cases are similar.

We first prove that a coordination game implies $|A| > |B|$, $|C| > |D|$, $A > 0$ and $C > 0$. First of all, when $a > c$ and $d > b$, A is positive, $A = \frac{1}{4}(a - c + d - b) > 0$. Then, because $A > 0$, the expression $|A| > |B|$ can be written as $A > |B|$. If $B > 0$, this expression can further be written as $A - B > 0$. This inequality is indeed satisfied from the condition $d - b > 0$ that defines a coordination game, i.e. $A - B = 2(d - b) > 0$. If $B < 0$, we need to check that $A + B > 0$, and this is obtained from the other coordination game condition $a - c > 0$, i.e. $A + B = 2(a - c) > 0$. The argument for C and D is analogous.

We next consider the viceversa—that the conditions on A, B, C and D imply a coordination game. Consider $B > 0$ without loss of generality. Because also $A > 0$, we can remove the absolute values in the condition $|A| > |B|$, which becomes $A - B > 0$. This implies $d > b$. We still need to show that $a > c$, which does not simply follow from the definition of A, $A = \frac{1}{4}(a - c + d - b) > 0$. Indeed, the definition of A only implies $a - c > -(d - b)$, but because we just proved that $d - b > 0$, this condition could also be satisfied with $a - c < 0$. However, if $a - c < 0$, we have $B = \frac{1}{4}(a - c - (d - b)) < 0$, which is in contradiction with our former assumption. The same considerations apply to C and D. □

2.2.2 Special cases of EWA

Here, we present the parameter restrictions at which EWA reduces to the learning algorithms it generalizes. We give more details than in Section 1.3.5. In Figure 2.1, we show three key parameters α, β and δ. The fourth parameter κ is not shown, but we specify its values in the following discussion. For convenience, we report here the EWA dynamical equations

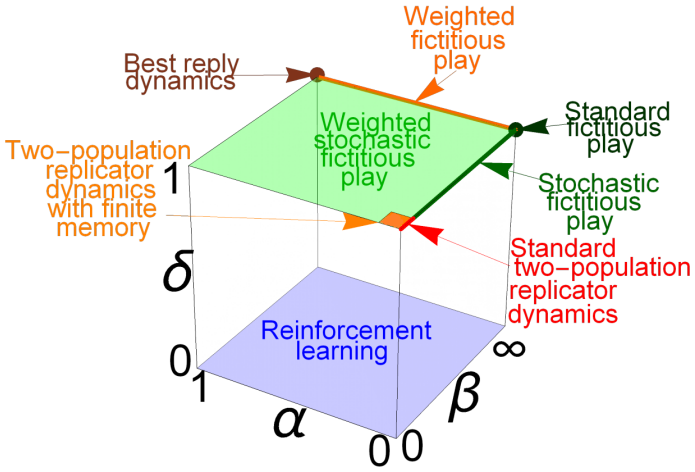


Figure 2.1: Learning algorithms generalized by EWA. In this figure we show three EWA parameters: the memory loss α, the payoff sensitivity β and the weight given to forgone payoffs δ. The fourth parameter κ is discussed in the main text. The case $\alpha = 0$, $\beta \to 0$, $\delta = 1$ corresponds to two-population replicator dynamics and is shown as a segment near the value $\beta = 0$; the analogous case with α very small but strictly positive, $\alpha \to 0$, is a finite-memory generalization of replicator dynamics and corresponds to the small square near the values $\alpha, \beta = 0$. This figure is an adapted version of the cube introduced in Ho et al. (2007), showing the parameter β in place of the parameter κ.

(1.29), (1.30), (1.31) for player Row (the equations for player Column are equivalent[20]):

$$\begin{aligned} x_i(t+1) &= \frac{e^{\beta Q_i^R(t+1)}}{\sum_j e^{\beta Q_j^R(t+1)}}, \\ Q_i^R(t+1) &= \frac{(1-\alpha)\mathcal{N}(t)Q_i^R(t)+(\delta+(1-\delta)I(i,s^R(t+1))\Pi^R(i,s^C(t+1))}{\mathcal{N}(t+1)}, \\ \mathcal{N}(t+1) &= (1-\alpha)(1-\kappa)\mathcal{N}(t)+1. \end{aligned} \tag{2.3}$$

The case $\alpha = 1$, $\beta = +\infty$, $\delta = 1$, for all values of $\kappa \in [0, 1]$, is best reply dynamics. Under best reply dynamics, each player only considers her opponent's last move (complete memory loss of previous performance, $\alpha = 1$), and plays her best reply to that move with certainty ($\beta = +\infty$). Playing the best reply generally requires fully taking into account the action that she did not play in the previous round of the game ($\delta = 1$). These parameter restrictions do not actually lead to the same best reply dynamics as the one described in Section 1.3.1 and used in Chapter 3. That version of best reply dynamics, indeed, assumes that players take turns in best responding to their opponent (alternating best reply dynamics). This is rather *simultaneous* best reply dynamics, because both players best respond at the same time. In 2×2 games, the only difference in asymptotic behavior between the two cases is in coordination and anticoordination games. In these games, if the players start at an action profile which is not a Nash equilibrium, under simultaneous best reply dynamics they "jump" back and forth between the non-equilibrium action profiles. Under alternating best reply dynamics, the would instead converge to equilibrium.

[20]This implies that we assume that the EWA parameters are the same for the two players. The principles of our analysis can however easily be extended to the case in which parameters are different.

The case $\alpha = 0$, $\beta = +\infty$, $\delta = 1$ (and $\kappa = 0$) corresponds to fictitious play (Section 1.3.3). This is the counterpart of best reply dynamics in terms of memory loss, in the sense that players determine their action by taking into account all past play equally. Relaxing the assumption of infinite memory, the case with $\alpha \in (0, 1)$ corresponds to *weighted fictitious play*, as more recent history of play carries greater weight in estimating the opponent's mixed strategy. If $\alpha = 0$ but $\beta < +\infty$, the players weigh all history of play equally, but they do not choose the action with highest attraction with certainty. This case corresponds to *stochastic fictitious play*. Finally, the combination of finite memory and finite intensity of choice, i.e. $\alpha \in (0, 1)$, $\beta < +\infty$, $\delta = 1$ (again with $\kappa = 0$), results in *weighted stochastic fictitious play*. The literature on the convergence properties of fictitious play and its variants has been surveyed in Section 1.3.3.

The limit $\beta \to 0$, with $\alpha = 0$, $\delta = 1$ and $\kappa > 0$, corresponds to *two-population replicator dynamics* (Section 1.3.4; for a derivation, see Appendix 2.A.1).[21] A full characterization of the convergence properties of this algorithm in 2×2 games is given in Hofbauer and Sigmund (1998) (Chapter 10). Two-population replicator dynamics converges to NE in all 2×2 games, except in cyclic ones, where it circles around the unique mixed strategy NE. Assuming instead that α is positive but small, $\alpha \to 0$ (s.t. the ratio α/β is finite) we obtain a generalized *two-population replicator dynamics with finite memory*, originally proposed by Sato and Crutchfield (2003). To the best of our knowledge, its dynamical properties have not yet been analyzed in 2×2 games.

Finally, when $\delta = 0$, EWA reduces to reinforcement learning. By only updating the propensities of the actions they took, players only consider the payoff they experienced, without thinking about what their opponent may play. If, additionally, $\alpha = 0$ and $\kappa = 1$, EWA reduces to the specific form of reinforcement learning in Eqs. (1.21)–(1.22) (Section 1.3.2).

2.2.3 Plan for the exploration of the parameter space

Our challenge is to characterize learning behavior in a 13-dimensional parameter space (the eight payoffs a, b, c, d, e, f, g, h; the four learning parameters α, β, δ and κ; and the specification of the learning rule, which can be deterministic or stochastic).

Due to non-linearity of EWA, it is impossible to obtain a closed-form characterization of the learning dynamics as a function of all parameter combinations. Therefore, we follow the modular strategy outlined in Table 2.2. In Section 2.3 we start with a baseline scenario

[21] Our derivation is different from Börgers and Sarin (1997) and Hopkins (2002) because these authors consider one of the versions of reinforcement learning proposed in Erev and Roth (1998), in which attractions map linearly to probabilities. We use instead a logit form. As a result, we get a different continuous time limit.

Case	Section	Payoffs	α	β	δ	κ	Rule
Baseline	2.3	$A = C$, $B = D$; $A = -C$, $B = -D$	–	–	$\delta = 1$	$\kappa = 1$	DET
Unconstrained payoffs	2.4.1	A, B, C, D	–	–	$\delta = 1$	$\kappa = 1$	DET
Belief learning	2.4.2	A, B, C, D	–	–	$\delta = 1$	–	DET
Reinforcement learning	2.4.3	a, b, c, d, e, f, g, h	$\alpha = 1$	–	–	$\kappa > 0$	DET
Stochastic learning	2.4.4	A, B, C, D	–	–	$\delta = 1$	$\kappa = 1$	STOCH

Table 2.2: Plan for the exploration of the parameter space. Where the parameters are not set to any value (–), it means that in principle we fully analyze the dynamics for each value that the parameter can take. We start from a baseline scenario in which only four parameters are not fixed. We then follow a modular strategy, as in each case we explore the effect of varying one or more parameters at a time. For example, the case that we name "unconstrained games" explores the effect of relaxing the constraints $A = \pm C$ and $B = \pm D$. In all cases with $\delta = 1$ the payoffs can be reduced to the combinations A, B, C, D, so we indicate these parameters only; for reinforcement learning, this is not possible and so we indicate instead all payoffs a, b, c, d, e, f, g, h. The last column considers whether the learning rule is stochastic (STOCH) or deterministic (DET). We argue that behavior under virtually any parameter combination can be understood based on the cases considered in this table.

in which dynamics is deterministic and only four parameters do not take a fixed value. We consider this scenario as the baseline because it allows a mathematical simplification that eases the analysis, and because it is a clear benchmark against which to compare other parameterizations. The parameters that are not fixed in the baseline scenario are the payoff combinations A and B (C and D are constrained to be either equal or of opposite sign than A and B), the memory loss α and the intensity of choice β. Here we are exploiting the fact that, for $\delta = 1$, only the payoff combinations A, B, C, D, and not all eight payoffs, determine the dynamics (see Section 2.3.1).

We then consider various extensions, exploring the effect of changing one or more additional parameters while holding the others constant. For example, in Section 2.4.1 we relax the constraint that $A = \pm C$ and $B = \pm D$, and consider the effect of different combinations of payoffs to the two players. In Section 2.4.2 we relax $\kappa = 1$ and fully explore the effect of changing κ in the interval between 0 and 1. (We also keep with the former extension of considering generic A, B, C and D.) The specific case $\kappa = 0$ corresponds to *belief learning* (Camerer and Ho, 1999), a class of learning algorithms that includes best reply dynamics and fictitious play, in which the players respond to their beliefs of what the opponents are going to play. In Section 2.4.3 we let δ vary in the interval between 0 and 1, fixing $\alpha = 1$ for analytical convenience. The specific case $\delta = 0$ corresponds to *reinforcement learning*. Finally, in Section 2.4.4 we analyze *stochastic learning*, relaxing the deterministic approximation explained in Section 1.3.5.

Finally, in Section 2.4.5 we discuss some additional parameter combinations. We argue that, although our analysis does not explicitly cover all the parameter space, we can qualitatively understand the learning behavior under virtually any choice of the parameters; behavior can be understood as a superposition of the cases outlined in Table 2.2.

2.2.4 Summary of the results

In this section, we give some intuition about the qualitative learning behavior in the baseline scenario and in the extensions. We also discuss the novelty of this project. Existing research mostly applies to the edges and vertexes of the cube in Figure 2.1. The main contribution of this project is to generalize these results by characterizing the learning behavior over a larger portion of the parameter space. This leads to interesting new insights.

In the baseline scenario (see Table 2.2), we find that three combinations of learning parameters/payoffs mostly determine the learning dynamics. The first combination is the ratio α/β. We find it sensible to call this combination of parameters "inattention", because when α/β is large the players disregard information about the received payoffs, either through high memory loss α or through small payoff sensitivity β. The other two combinations of parameters are the products AC and $|BD|$. We name AC a "coordination" parameter, as it is positive and larger than $|BD|$ in coordination and anticoordination games. When $A = C$, the parameter AC is always positive; when $A = -C$, it is always negative. We call $|BD|$ a "dominance" parameter, as $|BD| > AC$ in dominance-solvable games, in which for each player one action dominates the other (we take the absolute value of the product BD because its sign only determines which action profile is preferred, but does not change the properties of game).

Figure 2.2 graphically depicts a qualitative characterization of the learning behavior in the baseline scenario. The top panel is a visualization of the three-dimensional parameter space spanned by α/β, AC and $|BD|$. The grey plane cuts the parameter space orthogonally to the α/β axis; the projection is reported in the bottom panel. The diagonals in the projection distinguish between the different classes of games; the colors represent the typical learning behavior. The class of the game and the typical learning behavior are clearly related, but there is no exact mapping (this is illustrated by the shaded boundaries between different color areas; see below for examples).

In the green area of the parameter space, the two players converge to a unique mixed strategy profile in which they choose both actions with similar probabilities. We refer to this as a "unique mixed fixed point". For example, Row may choose s_1^R with probability 0.55 and Column may select s_1^C with probability 0.60. This situation occurs when α/β is high compared to AC or $|BD|$. This is intuitive: the players quickly forget or are not

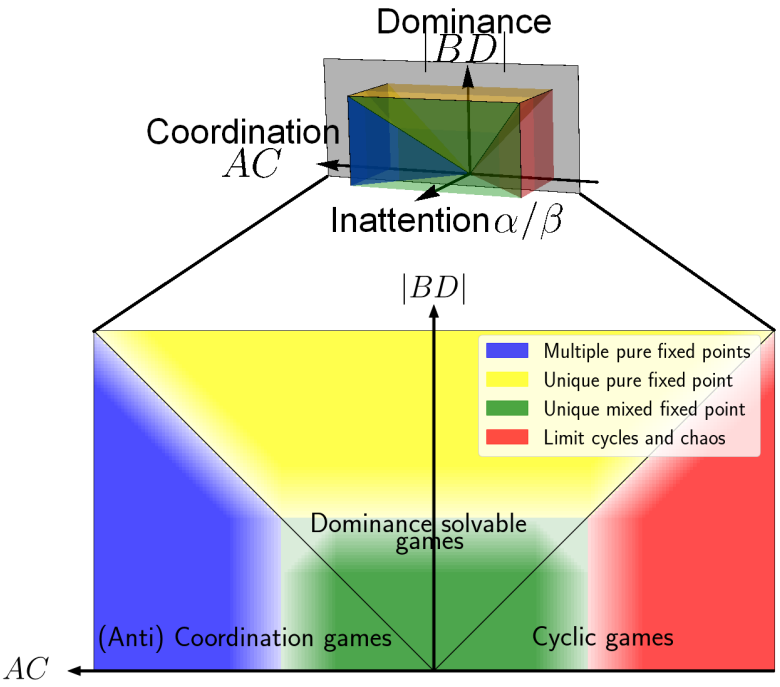


Figure 2.2: Qualitative characterization of the parameter space in the baseline scenario ($A = \pm C$, $B = \pm D$). The *inattention* (α/β) parameter is large when the players have short memory or disregard information about the payoffs. *Coordination* (AC) and *dominance* ($|BD|$) quantify properties of the payoff matrix and relate to well-known classes of 2×2 games. Inattention, coordination and dominance span a 3-dimensional parameter space, of which we provide a 2-dimensional cut with a plane orthogonal to the α/β axis. The relative signs and magnitudes of AC and $|BD|$ identify the relevant classes of 2×2 games; for example, given the constraint on the payoffs, $AC > 0$ and $AC > |BD|$ identifies coordination and anticoordination games (note positive values of AC are shown on the left). We give different colors to the areas of the parameter space with different learning dynamics. The boundaries between these areas are shaded as there is no clear-cut transition with the class of the game. This is different from simpler learning algorithms such as replicator dynamics and fictitious play, in which the class of the game alone and not the precise values of the payoffs determines the learning behavior.

sensitive to the payoffs, they do not really learn from playing the game and so end up just randomizing between their actions. Equivalently, for fixed α/β, if AC or $|BD|$ are small there are low *incentives* to learn. This means that players do not gain much for focusing on the action that yields higher payoff, so they just randomize between their actions. There is a scaling equivalence between increasing α/β and decreasing AC and $|BD|$.

In the yellow area, learning converges to a unique mixed strategy profile in which both players select one of their actions with high probability, and the other action with low probability. We refer to this situation as a "unique pure fixed point", because the behavior of the players is close to a pure action profile. This action profile corresponds to the pure Nash equilibrium of the dominance-solvable game that is described by the payoffs AC and $|BD|$.

When AC is positive and considerably larger than $|BD|$ and α/β (blue area), learning

converges to multiple fixed points, close to the two pure strategy NE of the corresponding coordination or anticoordination game. One of the two NE may be Pareto-inferior, in the sense that, by switching to the other NE, at least one player may be able to improve his payoff without reducing the payoff of the other player. If one of the two NE is Pareto-inferior, for sufficiently high α/β (or, equivalently, for values of AC that are not much larger than $|BD|$), the fixed point corresponding to the Pareto-inferior NE may disappear. The learning dynamics would then always end up in the Pareto-optimal equilibrium. This result is in contrast with simpler learning algorithms such as (two-population) replicator dynamics or standard fictitious play, in which pure strategy NE are always locally stable fixed points (Hofbauer and Sigmund, 1998). It also suggests that inattention is a method to achieve coordination on a Pareto-optimal NE, in a conceptually similar way to the theory of stochastic stability (Young, 1993; Gintis, 2000).

Finally, when AC is negative and considerably larger than $|BD|$ and α/β (red area), the players fail to converge to any fixed point. To illustrate the mechanism behind this instability, consider a Matching Pennies game in which the unique fixed point of the learning dynamics is the mixed strategy NE characterized by $x = 1/2$ and $y = 1/2$. Suppose that Row deviates and plays mixed strategy $x = 2/3$. As a result, on the following time step Column increases her probability to play s_2^C, say her mixed strategy becomes $y = 1/3$. When α/β is low, or the payoffs are high, the players over-react to the behavior of their opponent and the learning dynamics does not return to the NE. This may result in cycling behavior or, in some situations, chaotic dynamics. By changing the learning parameters we observe a bifurcation between stable and unstable dynamics. It is important to note that this bifurcation is only possible for finite memory and finite intensity of choice. This result generalizes Cheung and Friedman (1997), who showed the existence of a bifurcation in weighted stochastic fictitious play (see Section 1.3.3).

The discussion so far has been introduced to only apply to the baseline scenario, but most of the results above also apply to the extensions listed in Table 2.2. In Section 2.4.1, for example, we consider unconstrained payoffs, in the sense that C and D can take any value with respect to A and B. One interesting case is the one in which one player has larger payoff combinations than the other, say $C \gg A$ and $D \gg B$. Most of the discussion of the baseline scenario still applies. The only difference is that the player with highest payoffs converges to fixed points closer to the pure strategy profiles.

To extend the analysis to general values of κ (Section 2.4.2) it is enough to rescale β by a function of α and κ, so all the discussion above on the classes of the games and the combinations of payoffs still applies. The only difference is that long memory makes instability *less* likely (indeed, recall that infinite-memory fictitious play always converges

to NE), as opposed to the baseline scenario in which long memory makes instability more likely.

A situation in which the results of the baseline scenario do not apply as straightforwardly is $\delta < 1$ (Section 2.4.3). For small values of δ, the players never realize that they could unilaterally switch to another action and so earn higher payoffs. It is intuitive that this allows several *lock-in* fixed points. We provide a precise characterization of the stability of these fixed points. We exemplify this with a Prisoner Dilemma game.[22] If $\delta = 1$, cooperation is never a fixed point, because players always realize that defecting yields higher payoff if the other player cooperates. For $\delta < 1$, however, cooperation can be a locally stable fixed point. If the payoff for mutual cooperation – provided it is positive – is smaller by a factor $1/\delta$ than the payoff for unilateral defection, mutual cooperation is a stable fixed point. This result has some analogies with the concept of stochastic collusion (Macy and Flache, 2002) as a method to achieve a cooperation equilibrium.

The last extension that we consider is the stochastic version of EWA learning dynamics, in which players sample actions from their mixed strategy vector (Section 2.4.4). We show that the same qualitative properties as in the deterministic case hold. For example chaos, in the sense of exponential divergence of nearby trajectories, is robust to stochasticity. When the underlying dynamical system is unstable, the stochastic sequences of actions show a significant autocorrelation, while there is no statistically significant autocorrelation when the dynamical system reaches a fixed point. However, despite the positive autocorrelation, the sequences of actions are not perfectly predictable, arguing against the hypothesis that players will easily learn to identify cyclical patterns in play and so converge to a Nash equilibrium.

Finally, in Section 2.4.5, we discuss a few parameter and payoff combinations that have not been covered in the previous sections. We argue that the learning behavior in the parts of the parameter space that we have not explicitly explored can be understood based on our analysis.

2.3 Baseline scenario

We first analyze the asymptotic dynamics of EWA learning for the baseline scenario described in Table 2.2. In Section 2.3.1, we list three simplifications of the dynamical equations that considerably help the analysis (but are not valid under some of the extensions in Section 2.4). In Section 2.3.2 we analyze the existence and stability of fixed points. In Section 2.3.3 we simulate the learning dynamics in settings where all fixed points are unstable.

[22] In our framework, we do not consider repeated strategies such as tit-for-tat. Players are myopic and only learn based on previous time steps.

2.3.1 Simplifications

As a first simplification, we focus on the long-time outcome of learning, and assume that the experience $\mathcal{N}(t)$ takes its fixed-point value $\mathcal{N}^\star = 1/(1-(1-\alpha)(1-\kappa))$. This assumption is valid at long times as long as $(1-\alpha)(1-\kappa) < 1$,[23] but in practice, for most values of the parameters, $\mathcal{N}^\star$ is reached after a few time steps.

Substituting the above fixed point into (1.30), the update rule becomes

$$Q_i^\mu(t+1) = (1-\alpha)Q_i^\mu(t) + [1-(1-\alpha)(1-\kappa)]\left[\delta + (1-\delta)I(s_i^\mu, s^\mu(t+1))\right]\Pi^\mu(s_i^\mu, s^{-\mu}(t+1)). \tag{2.4}$$

Our second simplification is to take a deterministic limit of the learning dynamics. (The rationale for a deterministic limit was given in Section 1.3.5.) We write $\overline{\Pi_i^R}(y(t))$ for the expected payoff to player Row from playing pure strategy s_i^R at time t, given that player Column plays mixed strategy $y(t)$. For example, for $s_i^R = s_1^R$, the expected payoff is $\overline{\Pi_1^R}(y(t)) = ay(t) + b(1-y(t))$. Similarly, we write $\overline{\Pi_j^C}(x(t))$ for the expected payoff for player Column from playing action s_j^C, for a fixed mixed strategy $x(t)$ of player Row. The indicator function $I(s_i^\mu, s^\mu(t))$ can be replaced by the mixed strategy component, so for example $I(s_1^R, s^R(t)) = x(t)$.

It is possible to combine Eqs. (1.29) and (2.4) and to formulate a closed map for $x(t), y(t)$.

$$\begin{aligned} x(t+1) &= \frac{x(t)^{1-\alpha}e^{\beta\tilde{\kappa}[\delta+(1-\delta)x(t)]\overline{\Pi_1^R}(y(t))}}{x(t)^{1-\alpha}e^{\beta\tilde{\kappa}[\delta+(1-\delta)x(t)]\overline{\Pi_1^R}(y(t))} + (1-x(t))^{1-\alpha}e^{\beta\tilde{\kappa}[\delta+(1-\delta)(1-x(t))]\overline{\Pi_2^R}(y(t))}}, \\ y(t+1) &= \frac{y(t)^{1-\alpha}e^{\beta\tilde{\kappa}[\delta+(1-\delta)y(t)]\overline{\Pi_1^C}(x(t))}}{y(t)^{1-\alpha}e^{\beta\tilde{\kappa}[\delta+(1-\delta)y(t)]\overline{\Pi_1^C}(x(t))} + (1-y(t))^{1-\alpha}e^{\beta\tilde{\kappa}[\delta+(1-\delta)(1-y(t))]\overline{\Pi_2^C}(x(t))}}, \end{aligned} \tag{2.5}$$

where $\tilde{\kappa} = 1-(1-\alpha)(1-\kappa)$.

We can obtain a continuous-time version of Eq. (2.5) by taking the limit $\alpha \to 0$ and $\beta \to 0$, such that the ratio α/β is finite. Further details can be found in Appendix 2.A.1. When $\delta = 1$ and $\alpha = 0$, with $\beta \to 0$, EWA learning reduces to the standard form of the replicator dynamics. When $\alpha > 0$ (although small), EWA reduces to a generalized form of the replicator dynamics with finite memory (Sato and Crutchfield, 2003; Galla and Farmer, 2013).

Our third simplification simply derives from the baseline scenario assumption $\delta = 1$. It is possible to introduce a coordinate transformation that simplifies the dynamics, and helps to

[23] This restriction is always valid unless $\alpha = 0$ and $\kappa = 0$, as in standard and weighted fictitious play. However, it is possible to ex-post recover the convergence properties of these learning algorithms by taking the limit $\alpha \to 0$ in the stability analysis (see Section 2.4.2).

make the study of EWA analytically tractable. Specifically, we introduce the transformation

$$\begin{aligned}\tilde{x} &= -\frac{1}{2}\ln\left(\frac{1}{x}-1\right),\\ \tilde{y} &= -\frac{1}{2}\ln\left(\frac{1}{y}-1\right),\end{aligned} \tag{2.6}$$

only valid for x, y within the simplex, $x, y \in (0,1)$. In terms of the transformed coordinates (and assuming $\delta = 1$), the map (2.5) reads

$$\begin{aligned}\tilde{x}(t+1) &= (1-\alpha)\tilde{x}(t) + \beta[1-(1-\alpha)(1-\kappa)](A\tanh\tilde{y}(t)+B),\\ \tilde{y}(t+1) &= (1-\alpha)\tilde{y}(t) + \beta[1-(1-\alpha)(1-\kappa)](C\tanh\tilde{x}(t)+D),\end{aligned} \tag{2.7}$$

underlining that in the baseline scenario the game only enters through the four quantities A, B, C and D defined in Eq. (2.2). Mathematically, the transformation of coordinates is a diffeomorphism; it leaves properties of the dynamical system such as Jacobian or Lyapunov exponents unchanged (Ott, 2002).

The original coordinates are restricted to $x(t) \in (0,1)$ and $y(t) \in (0,1)$, the transformed coordinates instead take values on the entire real axis. Pure strategies $(x,y) \in \{(0,0),(0,1),(1,0),(1,1)\}$ in the original coordinates map to $(\tilde{x},\tilde{y}) \in \{(\pm\infty,\pm\infty)\}$ in the transformed coordinates, with 0 mapping to $-\infty$ and 1 mapping to $+\infty$ (but for these values the transformation is not valid).

2.3.2 Fixed point analysis

We now discuss existence and stability of fixed points. From now until the end of Section 2.3, we set $\kappa = 1$, in line with the parameter choices in the baseline scenario in Table 2.2.

As can be seen in Eq. (2.5), all pure strategy profiles are EWA fixed points. Indeed, if one replaces $x(t) = 0, 1$ on the right hand side in the first equation, $x(t+1)$ always takes the same value, and similarly $y(t)$. Intuitively, a pure strategy profile i, j corresponds to infinite propensities Q_i^R and Q_j^C, and finite changes in propensities (Eq. 2.3) have no effect. However, unless $\alpha = 0$, all pure strategy fixed points are unstable. (If $\alpha = 0$, only the Nash equilibria are stable pure strategy fixed points.) This is proved in Section 2.A.2. EWA also has one or three mixed strategy fixed points. In the following, we characterize existence and stability of the mixed strategy fixed points.

The location of the mixed strategy fixed points $(\tilde{x}^\star, \tilde{y}^\star)$ can be obtained from rearranging Eq. (2.7). The fixed points are the solutions to $\tilde{x}^\star = \Psi^R(\tilde{x}^\star)$ and $\tilde{y}^\star = \Psi^C(\tilde{y}^\star)$, where

$$\begin{aligned}\Psi^R(\tilde{x}^\star) &= \frac{\beta}{\alpha}\left[A\tanh\left(\frac{\beta}{\alpha}(C\tanh\tilde{x}^\star + D)\right) + B\right],\\ \Psi^C(\tilde{y}^\star) &= \frac{\beta}{\alpha}\left[C\tanh\left(\frac{\beta}{\alpha}(A\tanh\tilde{y}^\star + B)\right) + D\right].\end{aligned} \tag{2.8}$$

It is already possible to note that the EWA parameters α and β combine as the ratio α/β, our inattention parameter. Moreover, there is a scaling equivalence between increasing α/β or decreasing the payoff combinations A, B, C, D, as multiplying α/β by a constant and dividing the payoffs by the same constant leaves the fixed point equations unchanged.

Existence and stability in symmetric games. We now focus on the case of symmetric games: this implies $A = C$ and $B = D$,[24] and in turn $\Psi^R(\cdot) = \Psi^C(\cdot) = \Psi(\cdot)$. (Note that in this project symmetric games are just a special case to simplify the analysis, there is nothing special about their properties.) While $\tilde{x}^\star = \Psi(\tilde{x}^\star)$ and $\tilde{y}^\star = \Psi(\tilde{y}^\star)$ implies that $\tilde{x}^\star$ and $\tilde{y}^\star$ take the same values, the *pairs* $(\tilde{x}^\star, \tilde{y}^\star)$ are found by replacing the values in the original fixed point equations (2.7). Depending on the class of the game and learning parameters, there can be either one or three mixed strategy fixed points.

The Jacobian of the map in the transformed coordinates (obtained from Eq. (2.7)) is given by

$$J|_{\tilde{x}^\star,\tilde{y}^\star} = \begin{pmatrix} 1-\alpha & \frac{A\beta}{\cosh^2(\tilde{y}^\star)} \\ \frac{A\beta}{\cosh^2(\tilde{x}^\star)} & 1-\alpha \end{pmatrix}, \tag{2.9}$$

and its eigenvalues are

$$\lambda_\pm = 1 - \alpha \pm |A|\,\beta \frac{1}{\cosh(\tilde{x}^\star)\cosh(\tilde{y}^\star)}. \tag{2.10}$$

The fixed point is stable if $|\lambda_\pm| < 1$. After some algebra this results in the stability condition

$$\frac{\alpha}{\beta}\cosh(\tilde{x}^\star)\cosh(\tilde{y}^\star) - |A| \geq 0. \tag{2.11}$$

We can now analyze the location and stability of fixed points as we vary A and B, while holding $\alpha/\beta = 1$. Indeed, up to this point only the combinations $(\beta/\alpha)A$ and $(\beta/\alpha)B$ matter, so changing the value of α/β is equivalent to rescaling the payoffs. It is in general not possible to obtain a closed-form solution for $\tilde{x}^\star$. Therefore, we first explore the parameter space by solving Eq. (2.8) numerically, and then provide a mathematical analysis for a number of limiting cases in which it is possible to obtain a closed-form solution.

Figure 2.3 shows the properties of the fixed points as we vary A and B. The solid black lines in the figure separate the regions of anti-coordination games, dominance-solvable games and coordination games. The different colors are associated with qualitatively different learning dynamics. In the blue region, there exist multiple stable (mixed-strategy) fixed points. In the green/yellow area there is only one stable fixed point. Through a linear

[24] In a symmetric game the identity of the players does not matter, i.e. the payoff to player μ from playing action s_i^μ against action $s_j^{-\mu}$ does not depend on μ. In formula, this means that $\Pi^R(s_i^R, s_j^C) = \Pi^C(s_j^R, s_i^C)$, so $A = C$ and $B = D$.

interpolation the color gradient reflects the distance of the fixed point from the center of the strategy simplex: as the point in the (A, B) plane becomes more yellow, the fixed point becomes closer to a pure strategy profile.

Figure 2.4 illustrates the nature of the fixed points with a few typical examples.

Unique fixed point near pure strategy NE:
In case (a) of Figure 2.4 the payoff matrix describes a dominance-solvable game, in which actions s_2^R and s_2^C are strictly dominated by actions s_1^R and s_1^C. The fixed point is indicated by a dark circle, and is located at $(x^\star, y^\star) = (0.95, 0.95)$, very close to the unique pure strategy NE at $(1, 1)$ (solid triangle). The fixed point is stable. As discussed at the beginning of this section, all pure-strategy profiles are unstable fixed points (light circles).

Multiple stable fixed points near pure-strategy NE:
Cases (b) and (d) are examples of anticoordination and coordination games respectively. Each of the two games has three NE, as is indicated by the triangles. In each example one NE involves a mixed strategy, and the other two equilibria are pure-strategy NE. In both cases the values of A and B are such that there are three fixed points of EWA learning. For both examples there exist a *"central" fixed point*, located near the mixed-strategy NE and unstable under the learning dynamics (cyan circle near the centre of strategy space), and two stable *"lateral" fixed points*.

The important difference between the two cases is that in (b) both pure-strategy NE are also Pareto equilibria, whereas in (d) (s_1^R, s_1^C) is both a NE and a Pareto equilibrium. This generates the asymmetry between the $A > 0$ and $A < 0$ semiplanes.[25] When $A > 0$ and B gets larger, the payoff discrepancy between the Pareto-efficient NE and the Pareto-inefficient NE increases. The stable lateral fixed point closest to the Pareto-inefficient NE collides with the unstable central fixed point, generating a *fold bifurcation* in which both fixed points disappear. Effectively, positive memory loss and non-infinite payoff sensitivity prevent the learning dynamics from getting stuck in a local minimum, and help reaching the Pareto-efficient NE.

Unique fixed point away from pure-strategy NE:
Case (c) also corresponds to an anticoordination game, but the payoffs are smaller than in the previous example. As the payoffs are smaller, there are less incentives to learn: the only

[25] Note that the discrepancy between coordination and anticoordination games here is an artifact of the symmetry assumption $A = C$ and $B = D$. Indeed, an anticoordination game with payoff matrix $\begin{pmatrix} 0,0 & 4,4 \\ 3,3 & 0,0 \end{pmatrix}$ is asymmetric, but perfectly equivalent to case (d) in terms of Pareto-efficiency. See also footnote 19.

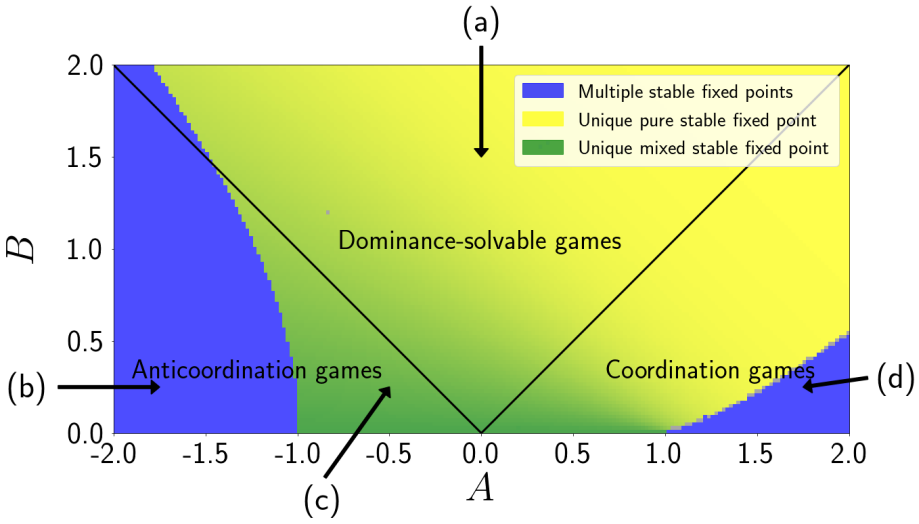


Figure 2.3: Quantitative characterization of the parameter space in the special case $A = C$, $B = D$ (symmetric games), for fixed $\alpha/\beta = 1$. The diagram shows the properties of the fixed points obtained from solving Eq. (2.8) numerically, and then checking the stability by using Eq. (2.11). We interpolate linearly between a "mixed fixed point" (green) and a "pure fixed point" (yellow), by making the color gradient proportional to the distance between the fixed point and the center of the strategy simplex. The annotations from (a) to (d) on the borders refer to specific games shown in Figure 2.4.

stable fixed point of EWA learning is near the centre of strategy space. This fixed point does not coincide with the mixed strategy NE, and the distance from the NE increases as either payoffs are lowered or the inattention parameter increased.

Mathematical analysis:

We proceed with the mathematical analysis for a number of specific cases. We first set $B = 0$. The boundaries between the blue and green areas in Fig. 2.3 is then found at $A = -1$ and $A = 1$. Mathematically, these boundaries mark the point at which the lateral fixed points cease to exist (they are present in the blue areas, but not in the green area). Calculating the slope of $\Psi(u)$ at $u = 0$, one shows that the lateral fixed points do not exist if

$$\frac{\beta}{\alpha}\,|A| \leq 1, \tag{2.12}$$

leading immediately to the boundaries $A = \pm 1$ in Figure 2.3 ($\alpha/\beta = 1$ in the figure).

When $\frac{\beta}{\alpha}\,|A| \to +\infty$, $\Psi(\tilde{x}^\star)$ approaches a step function equal to $-\frac{\beta}{\alpha}\,|A|$ in the negative domain and to $\frac{\beta}{\alpha}\,|A|$ in the positive domain, so the intersections with the $\tilde{x}^\star$ line occur precisely at $\tilde{x}^\star = 0$ and $\tilde{x}^\star = \pm\frac{\beta}{\alpha}\,|A|$. Recalling the mapping from the transformed coordinates to the original coordinates, these intersections correspond to $x = 0$, $x = 1/2$ and $x = 1$. By using the same argument for y, it is easy to see that the fixed points are the pure strategy NE of the coordination/anticoordination game and the mixed equilibrium in the center of the strategy space. In Figure 2.4, cases (b) and (d) approximate this situation.

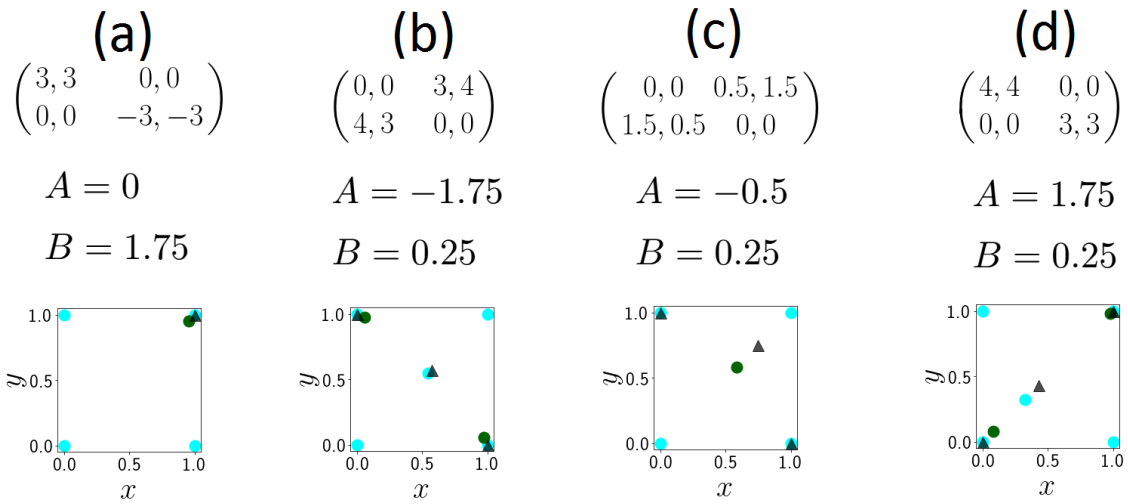


Figure 2.4: Payoff matrices and fixed point stability for games with certain values of A and B (Figure 2.3). The position of the fixed points of EWA and the NE are shown in the (x, y) plane. Green circles are stable fixed points; cyan circles are unstable fixed points; grey triangles are NE.

We now consider $B \neq 0$. If $\frac{\beta}{\alpha}|B| \to +\infty$ and $B \gg A$, $\Psi(\tilde{x}^\star)$ is completely flat and equal to $\Psi(0) = \frac{\beta}{\alpha}B$. This is also the position of the unique fixed point $\tilde{x}^\star$. As $\tilde{x}^\star \to \pm\infty$ (depending on the sign of B), $x \to 0, 1$ and the fixed point corresponds to the unique pure strategy NE. Case (a) in Figure 2.4 approximates this situation.

Stability is addressed in the following proposition.

Proposition 2. *In symmetric 2×2 games and with the learning parameters taking values as in the baseline scenario (Table 2.2) the following results hold:*

(i) if $B = 0$ and $\frac{\beta}{\alpha}|A| \leq 1$, the unique central fixed point is stable.

(ii) if $B = 0$ and $\frac{\beta}{\alpha}|A| \to 1^+$ or $\frac{\beta}{\alpha}|A| \to +\infty$, the central fixed point becomes unstable and the lateral fixed points are stable. In particular, at $\frac{\beta}{\alpha}|A| = 1$ a supercritical pitchfork bifurcation occurs.

(iii) if $\frac{\beta}{\alpha}|B| \to +\infty$ and $B \gg A$, the unique fixed point is stable.

The proof is in Appendix 2.A.3. In sum, in symmetric 2×2 games at least one fixed point is always stable, at least in the limiting cases covered in the proposition (but the numerical analysis above suggests that the results in the proposition also hold for intermediate values).

Existence and stability in asymmetric games. We focus on a specific type of asymmetric games in which the asymmetry only stems from the sign of the payoffs. These games are defined by the condition $\Pi^R(s_i^R, s_j^C) = -\Pi^C(s_j^R, s_i^C)$, which implies $A = -C$, $B = -D$. Note that this condition does not generally define zero-sum games, which are rather defined

by the equality $\Pi^R(s_i^R, s_j^C) = -\Pi^C(s_i^R, s_j^C)$.[26] Under this definition, if $B > A$ the game is dominance-solvable, but if $A > B$ we have a cyclic game.

Fixed points and stability:

As above, we first write down the conditions for the existence and stability of fixed points, and then study their properties as we vary the learning parameters and the payoffs.

When $A = -C$, Eqs. (2.8) have at most one solution, as the functions on the right hand side monotonically decrease. Moreover, if $B \neq 0$ we generally have $\tilde{x}^\star \neq \tilde{y}^\star$. The eigenvalues of the Jacobian (2.9) are complex and of the form

$$\lambda_\pm = 1 - \alpha \pm i \frac{\beta\,|A|}{\cosh(\tilde{x}^\star)\cosh(\tilde{y}^\star)}. \tag{2.13}$$

The stability condition is then:[27]

$$\frac{\beta}{\sqrt{2\alpha - \alpha^2}} \frac{|A|}{\cosh(\tilde{x}^\star)\cosh(\tilde{y}^\star)} \leq 1. \tag{2.14}$$

This stability condition is different from the one of symmetric games, in Eq. (2.10). Indeed, it is not only the ratio α/β that matters, but a more complicated function of these parameters. In general, increasing α or β has the same effect on stability as with the ratio α/β, but when taking the limit $\alpha, \beta \to 0$ (such that the ratio α/β is finite), the left hand side of the above equation goes to zero, and so the fixed point is always stable. This is consistent with replicator dynamics with finite memory always converging to a mixed strategy fixed point, which could however be arbitrarily far from a Nash equilibrium (see Appendix 2.A.1).

Examples of typical behaviour:

In Figure 2.5 we illustrate the different possible outcomes for asymmetric games, as we did for symmetric games in Fig. 2.3. Example (a) is a dominance-solvable game. The learning dynamics converges to a unique fixed point close to the pure strategy NE, analogously to case (a) in Figure 2.3. In case (b) we have instead a cyclic game, with relatively low values of the payoffs. As in symmetric games, low values of the payoffs imply that the fixed point at the center of strategy space — not necessarily corresponding to the NE — is stable. Case (c) is similar to case (b), but the payoffs are larger. Higher incentives make the players overreact to their opponent's actions, and this makes all fixed

[26] These asymmetric games and zero-sum games only correspond if $\Pi^R(s_i^R, s_j^C) = \Pi^C(s_i^R, s_j^C) = 0$ for $i \neq j$.

[27] Here we just find the condition under which $(\tilde{x}^\star, \tilde{y}^\star)$ — the only potentially stable fixed point — loses stability. It is possible to prove that the dynamical system undergoes a supercritical Hopf bifurcation (or Neimark-Sacker bifurcation) when the eigenvalues cross the unit circle. However, the proof involves calculating the so-called first Lyapunov coefficient, which requires a lot of algebra and does not provide any insight, so we do not provide a proof here. We instead use numerical simulations to show that the Hopf bifurcation is indeed supercritical.

points unstable. The learning dynamics gets trapped in limit cycles or, for some parameters, chaotic attractors, as we will show in Section 2.3.3.

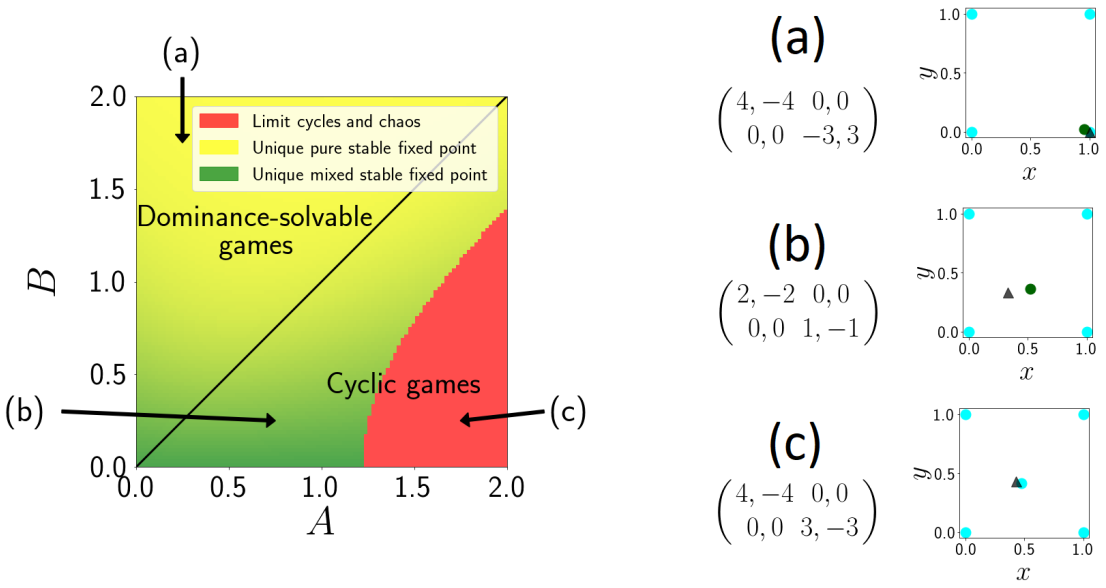


Figure 2.5: Quantitative characterization of the parameter space of asymmetric games in which $A = -C$ and $B = -D$, for $\alpha = \beta = 0.8$. This meaning of this figure is analogous to Figure 2.3. The central diagram corresponds to the $AC < 0$ semiplane in Figure 2.2. In the red portion of the parameter space no fixed points are stable, and the learning dynamics follows limit cycles or chaos.

Cyclic games – matching pennies:

We next focus on a specific example of cyclic games, Matching Pennies. This is a zero-sum game in which one player gains a coin, while the other player loses the coin (Osborne and Rubinstein, 1994). The resulting payoff matrix implies $B = D = 0$, $C = -A$. The learning dynamics have a unique fixed point at $(\tilde{x}^\star, \tilde{y}^\star) = (0, 0)$. Replacing in Eq. (2.14) we find that the fixed point is stable if

$$\frac{\beta}{\sqrt{2\alpha - \alpha^2}}|A| \leq 1. \tag{2.15}$$

For the values of α and β used in Figure 2.5, the fixed point becomes unstable for $A^\star = 1.224$. This corresponds the the boundary of the green and red areas for $B = 0$, at the bottom of Fig. 2.5.

Summing up, in asymmetric games defined by the constraint $A = -C$ and $B = -D$, there exists one stable fixed point unless $A > B$, in which case the fixed point may lose stability through a Hopf bifurcation.

2.3.3 Simulations of unstable dynamics

All analysis so far was about local stability of fixed points. We now simulate dynamics to assess global stability and to check which type of dynamics arise when all fixed points are unstable.

In symmetric games, dynamics always converge to one of the stable fixed points, except in one case. When β is large, α is small and $|A| \gg |B|$ (coordination or anticoordination game), it is possible to observe a stable limit cycle of period 2. In this cycle the players "jump" between the pure strategy profiles that are not NE of the coordination/anticoordination game. This is unsurprising, as these parameter restrictions make EWA closely resemble simultaneous best reply dynamics (see Section 2.2.2). As this dynamics is behaviorally unrealistic and not robust to stochasticity — it is enough that one player "trembles" and the dynamics converges to the NE — we ignore it for the rest of the analysis. It is just an artifact of the deterministic approximation.

In asymmetric games in which all EWA fixed points are unstable, we instead observe more behaviorally realistic strategic oscillations. To illustrate the nature of the unstable solutions, Figure 2.6 shows some examples of the learning dynamics for some values of α, β, A and B. In panels (a) to (c) we have $A = -C = 2$ and $B = D = 0$, while in panel (d) we consider $A = -C = -3.4$ and $B = -D = -2.5$.

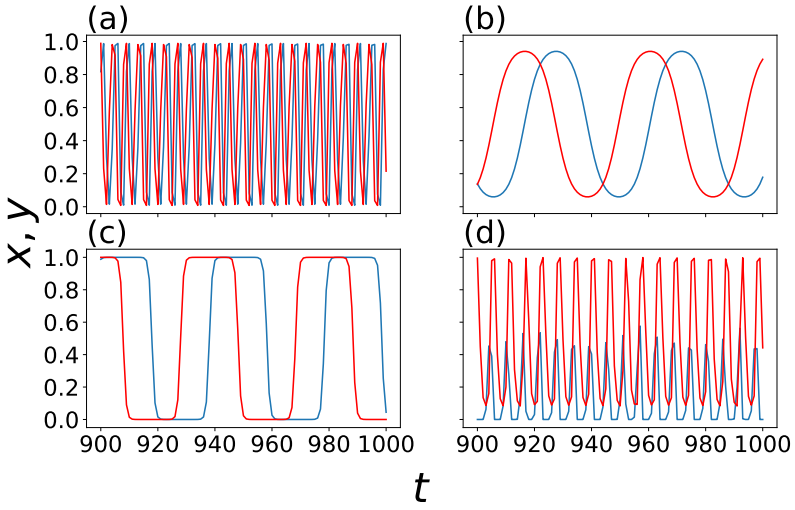


Figure 2.6: Time series of the probabilities x (in blue) and y (in red), for four different combinations of learning parameters and payoffs (detailed in the text). Cyclical and chaotic dynamics occur.

In panel (a), for $\alpha = 0.7$ and $\beta = 1$, the players frequently change their strategies, whereas in panel (b), for $\alpha = 0.01$ and $\beta = 0.1$, the dynamics is smoother. Note that the ratio α/β is very similar in the two cases, but nonetheless the dynamics is quite different. This is not in contradiction with the rest of the paper: only the fixed point behavior of

EWA is determined by the ratio α/β. In panel (c), where $\alpha = 0.01$ and $\beta = 0.5$ the players spend a lot of time playing mostly one action and then quickly switch to the other action (because they have long memory and high payoff sensitivity). Finally, in panel (d), we choose $B \neq 0$: this seems to yield the most irregular dynamics.

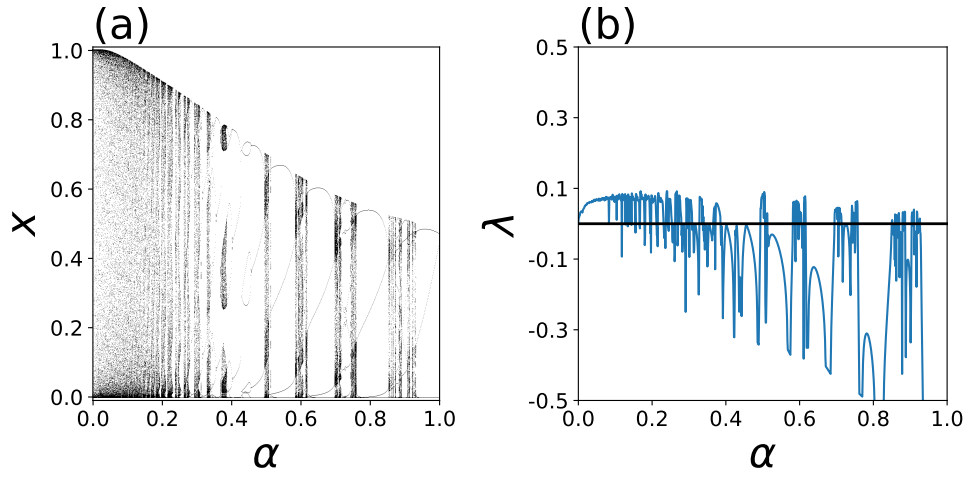


Figure 2.7: Bifurcation diagram and largest Lyapunov exponent λ as α is varied between 0 and 1. Cyclical and chaotic dynamics alternate, with chaos being more likely for small values of α.

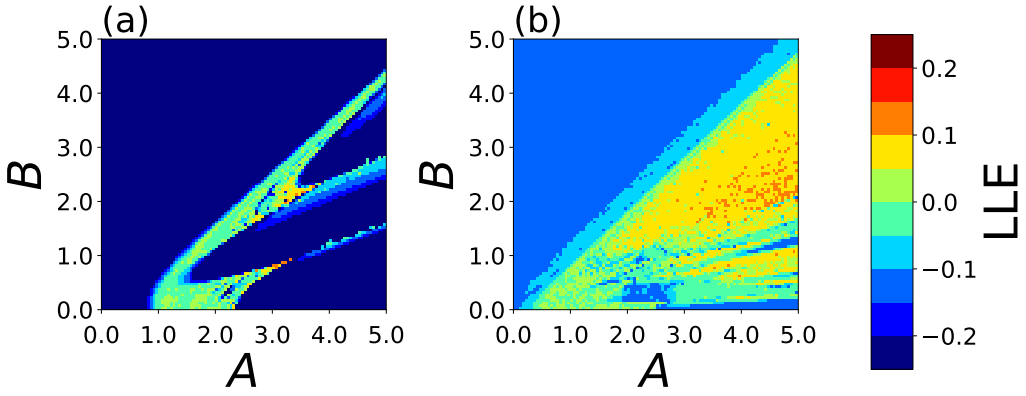


Figure 2.8: Largest Lyapunov exponent as a function of A and B in antisymmetric games ($C = -A, D = -B$). Colors from green to red indicate chaotic dynamics, while blue colors indicate convergence to a periodic attractor, which can be a limit cycle or a fixed point. In panel (b) players have longer memory.

Chaotic dynamics:

To check if the dynamics are chaotic or (quasi-)periodic, we consider a bifurcation diagram and calculate the Lyapunov exponents. In Fig. 2.7 we fix one payoff matrix (we use the example of Fig. 2.6(d), i.e. $A = -C = -3.4$ and $B = -D = -2.5$) and set the sensitivity of choice to $\beta = 1$. We then vary the memory-loss parameter α. All fixed points are unstable for any $\alpha \in [0, 1]$. In panel (a) we show the resulting bifurcation diagram. For each value of

α we plot the coordinates x the dynamics visits over the course of the trajectory, discarding an initial transient. When there are only a few values of x, e.g. for $\alpha \in [0.4, 0.5]$, the dynamics cycles between these values. When instead for a given value of α the dynamics visits significant portions of the phase space, as in $\alpha \in [0, 0.2]$, the dynamics is chaotic. This is confirmed in panel (b) where we plot the largest Lyapunov Exponent (LLE) λ; this exponent quantifies the exponential divergence of nearby trajectories (Ott, 2002), positive values indicate chaotic dynamics.

Figure 2.8 shows that chaos is more frequently observed if the players have long memory. Indeed, in panel (b) we set $\alpha = 0.01$, $\beta = 1$, while in panel (a) it is $\alpha = 0.7$. Chaos occupies a larger portion of the parameter space if one of the actions is dominant over the other, i.e. $B > 0$, as opposed to the case $B = 0$. The LLE is always negative if $B > A$, as the dynamics reaches a fixed point (consistently with the diagram depicted in Figure 2.5). The LLE is larger for intermediate values of the payoffs, i.e. for large A and B.

2.4 Extensions

We now consider the extensions to the baseline scenario (see Table 2.2). In Section 2.4.1 we consider games in which payoffs are not constrained by $A = \pm C$ and $B = \pm D$, so that the magnitude of the payoffs can be different to the two players. In Section 2.4.2 we consider values of the parameter $\kappa \in [0, 1)$ (in the case $\kappa = 0$ we recover belief learning). In Section 2.4.3 we consider $\delta \in [0, 1)$, recovering reinforcement learning for the case $\delta = 0$. Finally, in Section 2.4.4 we drop the simplification of deterministic learning and analyze the stochastic learning dynamics.

2.4.1 Unconstrained payoffs

From the point of view of learning, games in which $A \neq C$ and $B \neq D$ are widely similar to games in the same class for which the constraint $A = \pm C$ and $B = \pm D$ holds. For example, dominance-solvable games with unconstrained payoffs are widely similar to dominance-solvable games with constrained payoffs.

The most interesting difference occurs for games in which payoffs to one player are substantially larger than payoffs to the other player. Without loss of generality, consider the case in which payoffs to Column are much larger than payoffs to Row. In this case $D >> B$ and $C >> A$. Having larger payoffs, Column has strongest incentives to play better performing actions, and so he plays a mixed strategy closer to the pure strategies. We illustrate this with specific examples in Figure 2.9, where we also show the functions $\Psi^R(\tilde{x}^\star)$ and $\Psi^C(\tilde{y}^\star)$. In panels (a) and (b) player Row has smaller payoffs, and so lower incentives. As a result, x is always closer to the center of the strategy space than y. In

panel (c) we show a similar payoff matrix to case (b) in Figure 2.5, except that the large payoffs of player Column make the unique fixed point of (2.8) unstable.

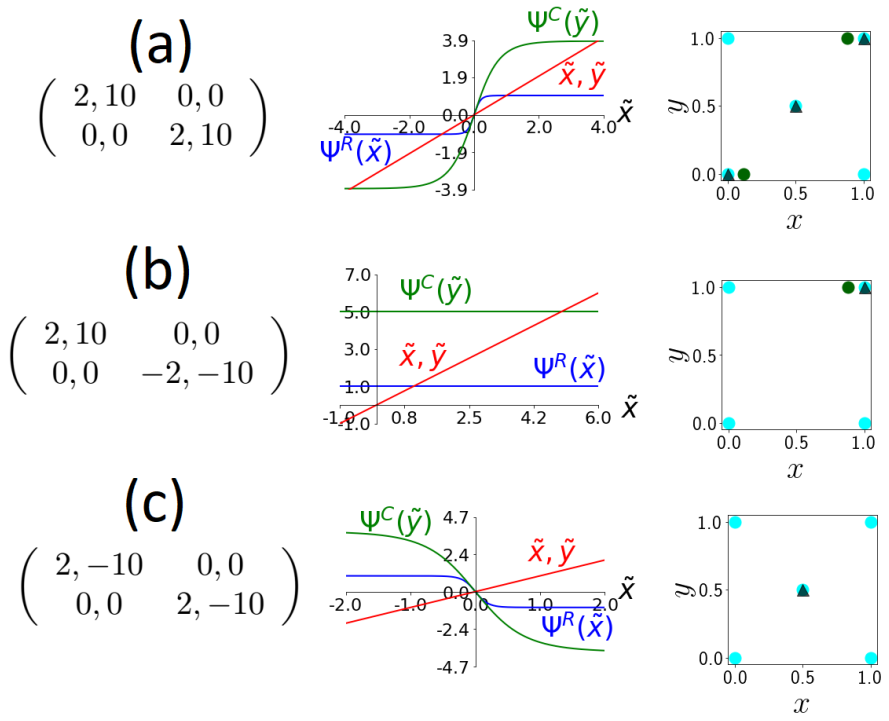


Figure 2.9: Examples of asymmetric games in which $B \neq D$ and $A \neq C$. These games are analogous to games with constrained payoffs ($A = \pm C$ and $B = \pm D$) in the same class, i.e coordination, dominance-solvable and cyclic games respectively for panels (a) to (c). The only difference is that the player with highest payoffs — and so strongest incentives — plays a mixed strategy closer to the pure strategies.

2.4.2 Belief learning

Choosing $\kappa < 1$ in Eqs. (2.5) and (2.7) is equivalent to rescaling the payoff sensitivity β as follows

$$\tilde{\beta} = \beta \left[1 - (1 - \alpha)(1 - \kappa)\right]. \tag{2.16}$$

As the quantity multiplying β is smaller than one for $\kappa < 1$, the effective payoff sensitivity is reduced, and so the inattention parameter becomes larger. Therefore, the learning dynamics is generally more stable for $\kappa < 1$, and convergence to a fixed point in the center of the strategy space occurs for a larger set of parameter combinations. All the analysis of the baseline scenario still applies.

In the belief learning case ($\kappa = 0$) the rescaled payoff sensitivity is $\tilde{\beta} = \beta\alpha$. This means that the coordinates of the fixed points do not depend on α, see Eqs.(2.8). One can show that the fixed points then correspond to the Quantal Response Equilibria (QRE) of the game. QRE were introduced by McKelvey and Palfrey (1995) to allow for boundedly

rational players, in particular to include the possibility that players make errors. Here the QRE $x^\star$ and $y^\star$ are given by the solutions to

$$\begin{aligned} \overline{\Pi_2^R}(y^\star) - \overline{\Pi_1^R}(y^\star) &= \frac{1}{\beta} \ln \frac{1-x^\star}{x^\star}, \\ \overline{\Pi_2^C}(x^\star) - \overline{\Pi_1^C}(x^\star) &= \frac{1}{\beta} \ln \frac{1-y^\star}{y^\star}. \end{aligned} \tag{2.17}$$

For small values of β the QRE are in the center of the strategy space, whereas increasing values of β bring the QRE closer to the NE. In the limit $\beta \to \infty$, the QRE coincide with the NE.

With $\kappa = 0$ the stability condition is

$$\frac{\beta\alpha}{\sqrt{2\alpha - \alpha^2}} A \leq 1. \tag{2.18}$$

Differently from Eq. (2.14), the derivative of the left hand side with respect to α is positive and so longer memory promotes stability. For general κ, the numerator in Eq. (2.18) is $\beta\left[1-(1-\alpha)(1-\kappa)\right]$, so the derivative is positive when $\alpha > \kappa$. The effect of memory on stability is thus not trivial: in the belief learning limit, long memory promotes stability, but when $\alpha > \kappa$ long memory promotes instability. To the best of our knowledge, we are the first to identify the role of memory on instability in this class of learning algorithms.

In the limit $\alpha \to 0$, the left hand side of Eq. (2.18) goes to zero, so stability is ensured for all parameter values. This way, we recover the well known result of Miyazawa (1961) and Monderer and Sela (1996), namely that in non-degenerate 2×2 games fictitious play would converge to the NE. For other values of β, we recover the results of Fudenberg and Kreps (1993) and Benaïm and Hirsch (1999), namely that in 2×2 games stochastic fictitious play would converge to the QRE.

2.4.3 Reinforcement learning

We now relax the constraint $\delta = 1$, and allow the players to give different weight to the actions that were and were not taken. For analytical tractability we assume that the players have perfect memory, $\alpha = 0$. We also assume $\kappa > 0$. (With $\alpha = 0$, β does not determine the existence and properties of the fixed points; so we could just set $\tilde{\beta} = \beta\left[1-(1-\alpha)(1-\kappa)\right] = 1$.) As we cannot use the coordinate transformation (2.6), we obtain the fixed points directly from Eq. (2.5).

By replacing the parameter restrictions in Eq. (2.5), it is possible to show that there are now potentially ten fixed points for a given value of δ. We give the expressions of all fixed points explicitly or implicitly in Appendix 2.A.4. Four fixed points are the pure strategy profiles, with (x, y) equal to $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$. In four additional fixed points

either x or y are 0 or 1, but not both, i.e. these fixed points are of the form $(0, y_1)$, $(1, y_2)$, $(x_1, 0)$ and $(x_2, 1)$. Finally, two fixed points have both x and y different from 0 and 1, i.e. (x_3, y_3) and (x_4, y_4). Only the pure strategy profiles are fixed points for all choices of model parameters; the other fixed points may or may not exist, depending on the choice of δ or of the payoffs.

In terms of stability, for each fixed point corresponding to the pure strategy profiles $(x, y) = \{(0,0), (0,1), (1,0), (1,1)\}$, we specify the two eigenvalues of the Jacobian at that fixed point:

$$\begin{aligned} (0,0) \quad &\rightarrow \quad \left(e^{\beta(b\delta-d)}, e^{\beta(f\delta-h)}\right), \\ (0,1) \quad &\rightarrow \quad \left(e^{\beta(a\delta-c)}, e^{\beta(h\delta-f)}\right), \\ (1,0) \quad &\rightarrow \quad \left(e^{\beta(d\delta-b)}, e^{\beta(e\delta-g)}\right), \\ (1,1) \quad &\rightarrow \quad \left(e^{\beta(c\delta-a)}, e^{\beta(g\delta-e)}\right). \end{aligned} \tag{2.19}$$

If we set $\delta = 1$, we get the same result mentioned at the beginning of Section 2.3.2 (see also Appendix 2.A.2), namely that only the pure strategy NE are stable. However, by taking $\delta < 1$ it is also possible to make the other pure strategy profiles potentially stable, by effectively reducing the "perceived" value of the payoffs at the NE (i.e., the players do not realize they could earn higher payoff if they unilaterally switched). We explain this with an example. Consider the strategy profile $(1, 1)$, and assume that $(0, 1)$ is a NE. This means that $c > a$, and so from Eq. (2.19) the first eigenvalue of $(1, 1)$ is greater than one for $\delta = 1$. So the strategy profile $(1, 1)$ is unstable. But if $\delta \neq 1$, the condition for the fixed point to be unstable is $c\delta - a > 0$, i.e. $c > a/\delta$. Therefore, provided $a > 0$, the payoffs at the NE must be larger by a factor $1/\delta$ if only the NE is to be a stable fixed point of EWA. In the other cases, non-NE can also be stable fixed points. Mathematically, this means that for $\delta < 1$ the dynamics can be stuck in local minima that are hard to justify in terms of rationality, as each player could potentially improve her payoff by switching action. However, as the players do not consider forgone payoffs, they cannot realize this, and keep playing the same action.

Explicit examples are given in Figure 2.10. The axes x and y give the positions of the fixed points for a specific value of δ, and the vertical axis δ shows how x and y vary with this parameter. When the lines are blue it means that the fixed point is stable, and when red the fixed point is unstable. Dashed green lines represent NE. When a NE coincides with a fixed point the lines are shown blue or red with green dashes.

In panel (a) the game is dominance-solvable, with a unique NE at $(x, y) = (0, 0)$. This NE is stable for all values of δ, but the pure strategy profiles $(1, 0)$ and $(0, 1)$ are also stable

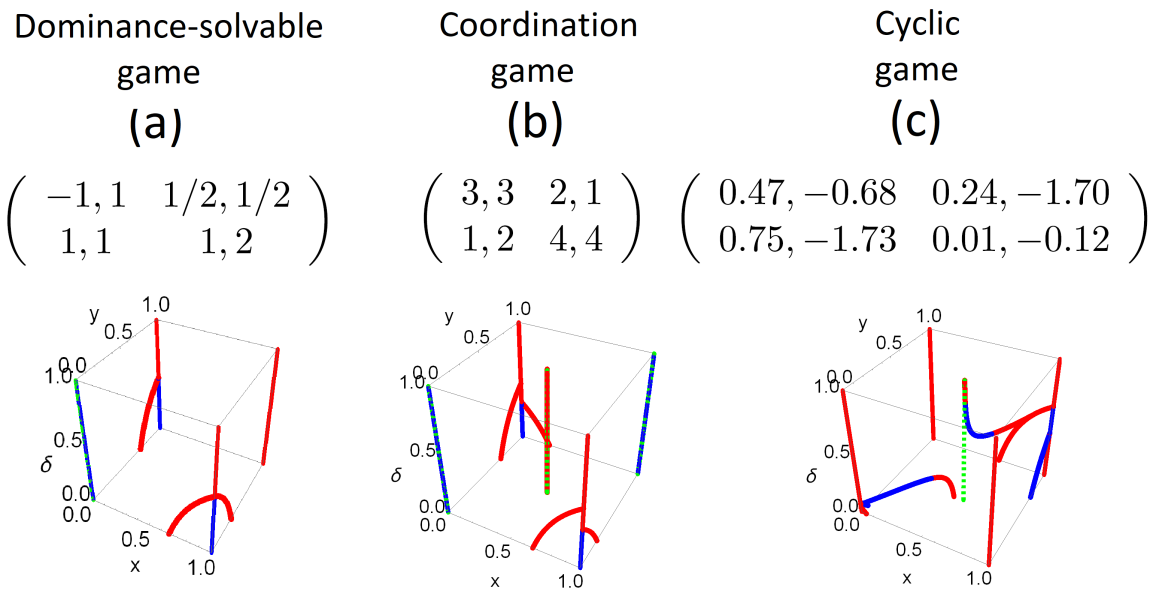


Figure 2.10: Bifurcation diagram showing the fixed points (x, y) as the δ parameter is varied between 0 and 1. Blue (red) lines represent stable (unstable) fixed points. Dashed green lines represent NE. Lower values of δ increase the likelihood to have stable fixed points that do not coincide with the NE.

for $\delta \in [0, 1/2]$. Up to some modifications of the payoffs this game can be interpreted as a Prisoner Dilemma game, in which for small values of δ cooperation can be sustained. The other solutions of the type $(0, y)$ or $(x, 0)$ — on the faces of the cube — are always unstable. Moreover, due to the specific values of the payoffs, the solution $(x_2, 1)$ does not exist. The situation is similar in panel (b), except that the payoff matrix describes a coordination game with two pure strategy NE. The other two pure strategy profiles are stable for $\delta < 1/3$, as can be calculated from Eq. (2.19).

Finally, case (c) is a cyclic game with the maximal number of fixed points, as all solutions exist. When $\delta = 0$ both fixed points $(0, y_1)$ and $(1, y_2)$ are stable; as δ is increased, the solution of the type (x_3, y_3) or (x_4, y_4) with $x, y < 0.5$ becomes stable. As δ is further increased the pure strategy profile $(1, 1)$ is stable, and finally it is the solution of the type (x_3, y_3) or (x_4, y_4) with $x, y > 0.5$ that becomes stable. For $\delta > 0.82$ all solutions are unstable, and the learning dynamics does not converge to any fixed point. Note that in this game no stable fixed point corresponds to the NE, and can be arbitrarily far.

2.4.4 Stochastic learning

When playing a game, except for very specific experimental arrangements (Conlisk, 1993a), real players update their strategies after observing a single action by their opponent, and so they do not not know her mixed strategy vector. This questions whether the analysis of

the deterministic dynamics so far provides robust conclusions. In this section we show that it does.

It is beyond the scope of this paper to systematically study the effect of noise on the learning dynamics. We refer the reader to Galla (2009) for a study on the effect of noise on learning, and to Crutchfield et al. (1982) for a more general discussion on the effect of noise on dynamical systems. We note two properties of this type of sampling noise.

First, noise is less likely to influence the dynamics if the players play strategies close to the pure strategies. In this case, the action they play will most likely coincide with the closest pure strategy. Consider for example a fixed point with $x, y = 0.95$, so that when both players choose action s_1 the probabilities x and y are fixed. The players only choose action s_2 on average $1/20$ of the times, and then return to the fixed point if it is stable. If instead the fixed point is in the center of the strategy space, both actions can be played with similar probability, and this leads to more irregular dynamics.

The non-linearity introduced by the logit function (1.29) also matters. The effect of an unlikely action — say action s_2 in the example above — on the strategy vector is weaker if players are close to the pure strategies. For example, if both $x, y = 0.95$ and the players choose s_2, for most parameter settings the dynamics deviates by very little (say, $\Delta x, \Delta y = 0.03$) from the fixed point. If instead the players were closer to the center of the strategy space, the deviation would be more substantial.

Given the remarks above, when the deterministic dynamics converges to a stable fixed point close to the pure strategies it is likely to stay there, with some occasional fluctuation. If instead there is a unique stable fixed point in the center of the strategy space, the fluctuations can be substantial. Simulations (not reported) confirm the conjectures above.

In Figure 2.11 we report instead examples where there are no stable fixed points, i.e., when the corresponding deterministic dynamics is cyclical or chaotic. The top panels represent stochastic learning, while the bottom panels show the corresponding deterministic learning dynamics. In all cases the payoff combinations are $A = -C = -3.4$ and $B = -D = -2.5$, and the memory loss is $\alpha = 0.2$. In the left panels the payoff sensitivity is $\beta = 1$ and so the deterministic dynamics is chaotic. The stochastic dynamics is very similar. This is mostly due to the chaotic attractor being close to the pure strategies, where the effect of noise is weak. In the right panels it is $\beta = 0.1$, so the deterministic dynamics reaches a fixed point. In the stochastic case, fluctuations induced by noise are much stronger than in the case of chaotic dynamics (panel (a)).

In Figure 2.11 we have shown time series of mixed strategies. However, real players cannot observe mixed strategies, and observe instead sequences of actions. We show such sequences of actions in Figure 2.12, where panel (a) corresponds to panel (a) in Figure

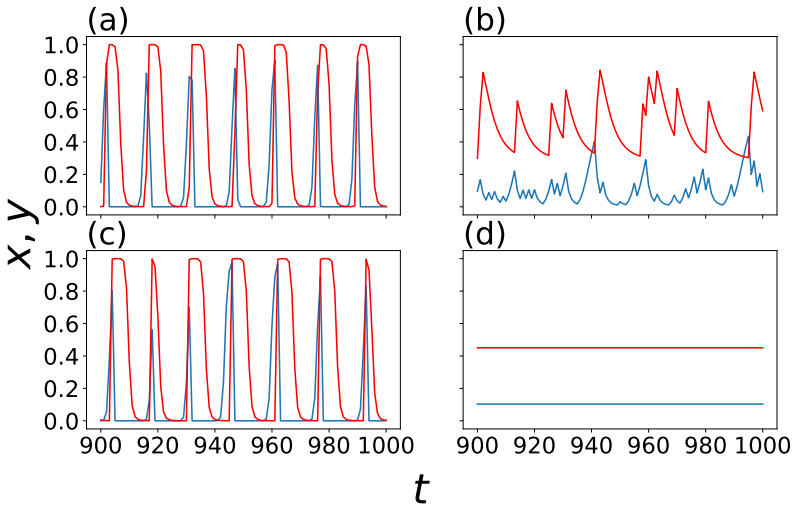


Figure 2.11: Time series of the probabilities x (in blue) and y (in red) of the learning dynamics in a cyclic game. Top panels represent stochastic learning, bottom panels show the corresponding deterministic learning. In panel (c) the deterministic dynamics converges to a chaotic attractor, while in panel (d) it reaches a fixed point.

2.11, and panel (b) corresponds to panel (b) in the same figure. The two sequences are qualitatively different, with the sequence in panel (a) showing much stronger persistence – that is, if one player has played a specific action, they are likely to play the same action in the next round.

To characterize this in more detail, we show the autocorrelation function of played actions in panel (c), for the two sequences in panels (a) and (b). The blue dashed lines show the region in which $|C(\tau)| < 0.1$: if the autocorrelation falls in this band the null hypothesis of no autocorrelation cannot be rejected at a 5% significance level. The sequence of actions in panel (b) shows no significant autocorrelation, while the sequence in panel (a) remains correlated up to $\Delta t \approx 3$. However, it should be noted that the sequence (a) is not at all perfectly predictable. Indeed, the number of times a certain action is played before switching to the other action varies after every switch.

Finally, in Figure 2.13 we show the bifurcation diagram and largest Lyapunov exponent as a function of α for stochastic learning. This figure is similar to Figure 2.7, consistently with theoretical studies on the effect of noise on dynamical systems (Crutchfield et al., 1982). The figure shows that chaos is robust to noise, the LLE is positive for $\alpha \in [0, 0.3]$. For $\alpha > 0.6$ the dynamics only visits a few points, as can be seen in panel (a). This is because the players have short memories and so only a few different histories of actions played are possible. In the extreme case of no memory, $\alpha = 1$, each player will "jump" between two points, corresponding to the two actions that her opponent may take at any time step. Indeed, in Figure 2.7(a) for $\alpha = 1$ the dynamics only visits two points ($x = 0$ and

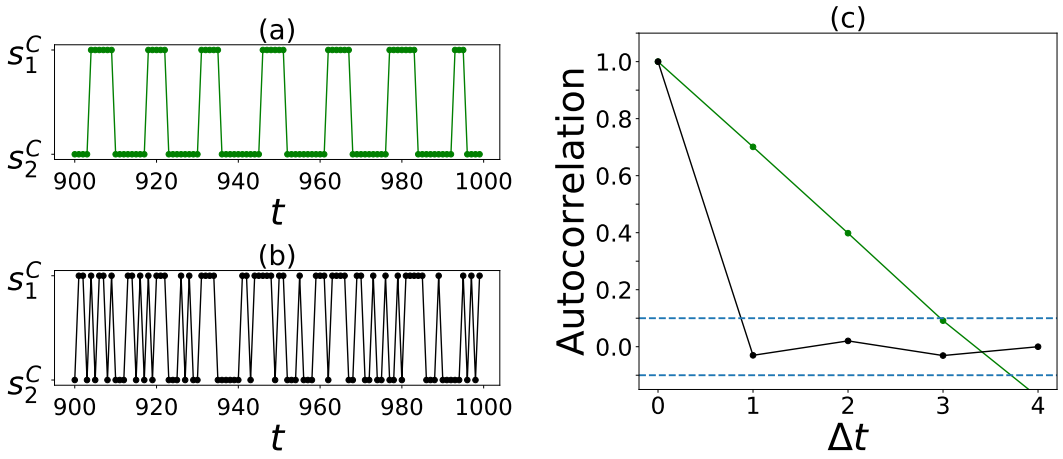


Figure 2.12: Left panels: sequences of actions taken by player Column. Panels (a)-(b) correspond to panels (a)-(b) in Figure 2.11. Right panel: Autocorrelation function of the actions of player Column, for both sequences represented in the left panel. When the learning dynamics reaches a fixed point in the deterministic limit, the autocorrelation decays instantaneously.

$x \approx 0.85$). This effect is absent in the deterministic dynamics, because the players choose distributions of actions.

2.4.5 Other parameter combinations

Here we cover some parameter and payoff combinations that have not been considered previously, and show that the results of our analysis can be directly applied to understand the learning behavior in these cases. We argue that in general, although we do not explicitly cover all the parameter space, all outcomes may be understood based on our analysis.

For example, consider the following game

$$\begin{pmatrix} 1,6 & 5,-2 \\ 3,2 & 1,-2 \end{pmatrix}. \tag{2.20}$$

Assume that $\delta = 1$, that the dynamics is deterministic, and consider any value of α, β and κ. What dynamics can we expect? The payoff combinations are $A = -1.5$, $B = 0.5$, $C = 1$ and $D = 3$. The payoffs do not satisfy any constraint of the type $A = \pm C$, $B = \pm D$. Differently from Section 2.4.1, moreover, the payoffs to one player are not simply a rescaled version of the payoffs to the other player, so that the magnitude of the payoffs is not the only difference with respect to the baseline scenario.

Nevertheless, we can qualitatively understand the outcome of learning based on our analysis. Because A and C have different signs, the functions $\Psi^R(\cdot)$ and $\Psi^C(\cdot)$ in Eq. (2.8) monotonically decrease, so there can only be one fixed point in the interior of the simplex. The game is dominance solvable and, although both B and D have the same

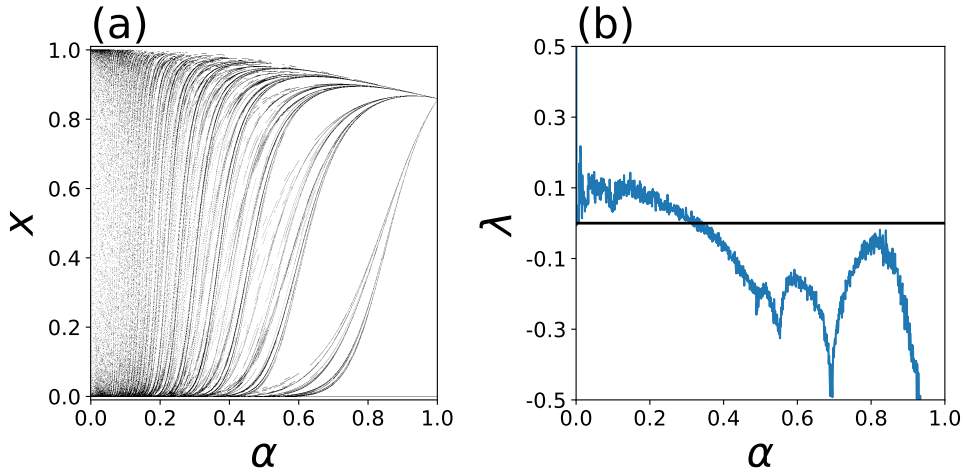


Figure 2.13: Bifurcation diagram and largest Lyapunov exponent λ as α is varied between 0 and 1. This figure is equivalent to Figure 2.7, except that here we consider stochastic learning. Chaos is robust to noise for small values of α.

sign, the situation is similar to the upper left corner of the diagram in Figure 2.5. If $\alpha/\tilde{\beta} = \alpha/\{\beta\,[1-(1-\alpha)(1-\kappa)]\}$ is small, the fixed point is close to the unique pure NE of the game, located at (s_2^R, s_1^C). If instead $\alpha/\tilde{\beta}$ is large, the fixed point is located in the center of the strategy simplex. Because $|D| > |B|$, finally, player Column always plays a strategy closer to the pure equilibrium than player Row, in line with the analysis in Section 2.4.1.

Relax now the assumption $\delta = 1$. Because we do not want to constrain α to take the value $\alpha = 0$, the analysis in Section 2.4.3 does not directly apply. However, combining the effects of δ and $\alpha/\tilde{\beta}$ is straightforward. Positive values of $\alpha/\tilde{\beta}$ push all fixed points towards the center of the simplex, so "lock-in" fixed points of the type discussed in Section 2.4.3 become less likely.

This is confirmed by simulating the EWA equations on the game in Eq. (2.20), fixing $\kappa = 0.5$, $\beta = 0.5$, and varying δ and α. We consider five values $\delta = 0.00, 0.25, 0.50, 0.75, 1.00$, and several values of $\alpha \in [0, 0.5]$. We simulate the dynamics and record the x variables in the final time steps, obtaining the bifurcation diagram in Figure 2.14. Consider the left panel, showing the deterministic dynamics. As $\alpha/\tilde{\beta}$ becomes larger, the fixed point moves towards the center of the simplex, in line with the analysis in Section 2.3. This is true for all values of δ. However, for $\delta = 0, 0.2$, an additional fixed point at $x = 1$ exists. This is one of the "lock-in fixed points" described in Section 2.4.3, and it only exists for sufficiently small values of $\alpha/\tilde{\beta}$.

Finally, in Section 2.4.4 we have discussed robustness of our results to stochasticity, but we have fixed κ and δ to $\kappa = 1$ and $\delta = 1$ respectively. The right panel of Figure 2.14 shows that robustness to stochasticity holds for other parameter values, too. (The

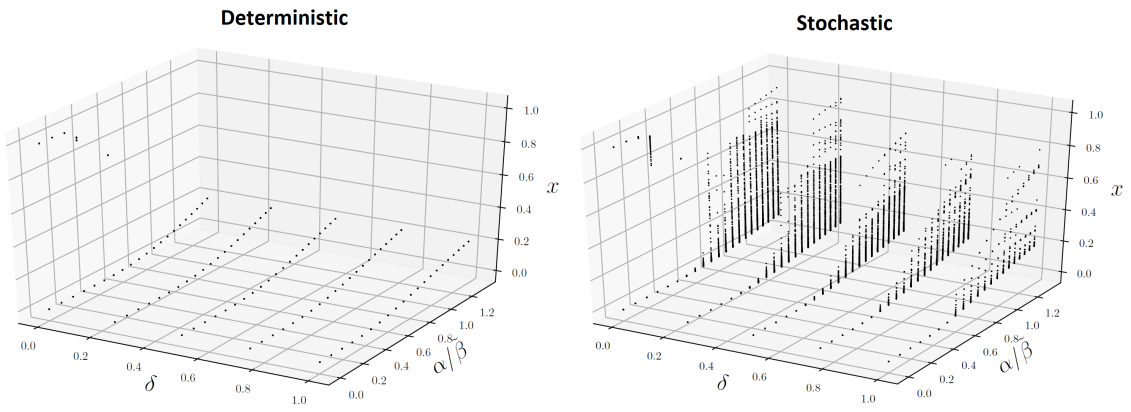


Figure 2.14: Bifurcation diagram obtained by simulating the EWA equations for some values of the parameters δ and $\alpha/\tilde{\beta}$ (we only show x, the mixed strategy of player Row). Fixed points with $x = 1$ are only possible for $\delta = 0, 0.2$ and for small values of $\alpha/\tilde{\beta}$.

density of points in the bifurcation diagram indicates that the value of x is often close to the deterministic one, although occasionally is it larger.)

While in this section we have not claimed that we can explicitly fully characterize the parameter space, we have shown a game and parameter combinations that had not yet been analyzed, but whose behavior could be qualitatively understood from the previous analysis.

2.A Appendix

2.A.1 Replicator dynamics with finite memory

Here we show that the EWA equations (2.5) have a continuous time limit that corresponds to a generalized version of replicator dynamics having finite memory, instead of infinite memory as in the standard case. We present an alternative derivation with respect to previous papers. Sato and Crutchfield (2003) assume that the evolution of the attractions takes place at a different timescale than the evolution of the probabilities, and Galla and Farmer (2013, Supplementary Information, Section II) use a Lagrange multiplier method.

Here we simply start from Eq. (2.5) and take the limit $\alpha \to 0$, $\beta \to 0$, such that the ratio α/β is finite. In this limit $\tilde{\kappa} = \kappa$, and we set $\kappa = 1$ without loss of generality.[28] We also only perform the calculations for $x(t)$, as the calculations for $y(t)$ are identical. For notational simplicity, we denote here $x(t)$ by x_t, and $y(t)$ by y_t. Taking logs in Eq. (2.5) we get

$$\ln x_{t+1} = (1-\alpha)\ln x_t + \beta[\delta + (1-\delta)x_t]\overline{\Pi_1^R}(y_t) - \ln\left(x_t^{1-\alpha} e^{\beta[\delta+(1-\delta)x_t]\overline{\Pi_1^R}(y_t)} + (1-x_t)^{1-\alpha} e^{\beta[\delta+(1-\delta)(1-x_t)]\overline{\Pi_2^R}(y_t)}\right). \quad (2.21)$$

[28] As the case $\kappa = 0$ is excluded from the condition on the steady state of the experience, κ is just a constant that multiplies β.

The logarithm of the denominator can be greatly simplified by taking the limit $\alpha \to 0$, $\beta \to 0$. In this limit

$$x_t^{1-\alpha} = x_t x_t^{-\alpha} = x_t e^{\ln x_t^{-\alpha}} = x_t e^{-\alpha \ln x_t} \approx x_t(1 - \alpha \ln x_t) \tag{2.22}$$

and

$$e^{\beta[\delta+(1-\delta)x_t]\overline{\Pi_1^R}(y_t)} \approx \left(1 + \beta[\delta + (1-\delta)x_t]\overline{\Pi_1^R}(y_t)\right). \tag{2.23}$$

By ignoring terms of $O(\alpha^2)$ (or equivalently $O(\beta^2)$, as the ratio α/β is finite) we can then write

$$\begin{aligned}
&\ln\left(x_t^{1-\alpha} e^{\beta[\delta+(1-\delta)x_t]\overline{\Pi_1^R}(y_t)} + (1-x_t)^{1-\alpha} e^{\beta[\delta+(1-\delta)(1-x_t)]\overline{\Pi_2^R}(y_t)}\right) \approx \\
&\qquad \ln\Big(x_t\left(1 - \alpha \ln x_t + \beta[\delta + (1-\delta)x_t]\overline{\Pi_1^R}(y_t)\right) + \\
&\qquad (1-x_t)\left(1 - \alpha \ln(1-x_t) + \beta[\delta + (1-\delta)(1-x_t)]\overline{\Pi_2^R}(y_t)\right)\Big) = \\
&\qquad \ln\Big(1 + x_t\left(-\alpha \ln x_t + \beta[\delta + (1-\delta)x_t]\overline{\Pi_1^R}(y_t)\right) + \\
&\qquad (1-x_t)\left(-\alpha \ln(1-x_t) + \beta[\delta + (1-\delta)(1-x_t)]\overline{\Pi_2^R}(y_t)\right)\Big) \approx \\
&x_t\left(-\alpha \ln x_t + \beta[\delta + (1-\delta)x_t]\overline{\Pi_1^R}(y_t)\right) + (1-x_t)\left(-\alpha \ln(1-x_t) + \beta[\delta + (1-\delta)(1-x_t)]\overline{\Pi_2^R}(y_t)\right).
\end{aligned} \tag{2.24}$$

Replacing this in Eq. (2.21) and rearranging gives

$$\begin{aligned}
\ln x_{t+1} - \ln x_t = \beta\Big(&[\delta + (1-\delta)x_t]\overline{\Pi_1^R}(y_t) - x_t[\delta + (1-\delta)x_t]\overline{\Pi_1^R}(y_t) - \\
&(1-x_t)[\delta + (1-\delta)(1-x_t)]\overline{\Pi_2^R}(y_t)\Big) - \alpha\left(\ln x_t - x_t \ln x_t - (1-x_t)\ln(1-x_t)\right).
\end{aligned} \tag{2.25}$$

It is possible to divide everything by β and rescale time so that one time unit is given by β. Then in the limit $\beta \to 0$ the left hand side of the above equation is

$$\lim_{\beta \to 0} \frac{\ln x_{t+\beta} - \ln x_t}{\beta} = \frac{\dot{x}}{x} \tag{2.26}$$

and the learning dynamics for x can be written in continuous time as

$$\begin{aligned}
\frac{\dot{x}}{x} = &[\delta + (1-\delta)x]\overline{\Pi_1^R}(y) - x[\delta + (1-\delta)x]\overline{\Pi_1^R}(y) - \\
&(1-x)[\delta + (1-\delta)(1-x)]\overline{\Pi_2^R}(y) - \frac{\alpha}{\beta}\left(\ln x - x \ln x - (1-x)\ln(1-x)\right).
\end{aligned} \tag{2.27}$$

This is in general the continuous time approximation of the EWA dynamics in Eq. (2.5). In the case $\delta = 1$, replacing the expressions for $\overline{\Pi_1^R}$ and $\overline{\Pi_2^R}$, we get

$$\frac{\dot{x}}{x} = ay + b(1-y) - (axy + bx(1-y) + c(1-x)y + d(1-x)(1-y)) - \frac{\alpha}{\beta}\left(\ln x - H(x)\right), \tag{2.28}$$

where $H(x) = x \ln x + (1-x)\ln(1-x)$ is the information entropy of mixed strategy $(x, 1-x)$. This is the dynamics analyzed in Sato and Crutchfield (2003). If $\alpha = 0$, i.e. with infinite memory, the above equation reduces to the standard form of two-population replicator dynamics (Hofbauer and Sigmund, 1998).

It is useful to analyze the stability of Eq. (2.28) in cyclic games. We rewrite the replicator dynamics (2.28) in terms of A, B, C and D, factor a $1-x$ term and write the corresponding equation for y:

$$\begin{aligned} \dot{x} &= x(1-x)\left(4Ay + 2(B-A) + \frac{\alpha}{\beta}(\ln(1-x) - \ln x)\right), \\ \dot{y} &= y(1-y)\left(4Cx + 2(D-C) + \frac{\alpha}{\beta}(\ln(1-y) - \ln y)\right). \end{aligned} \tag{2.29}$$

In line with the analysis in Section 2.4.1, we focus on the specific case in which $B = D = 0$ and $C = -A$, i.e. Matching Pennies. In this case the fixed points of the replicator dynamics are $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ and $(1/2, 1/2)$. The fixed points $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$ are always unstable, for any value of α. The eigenvalues for the fixed point $(1/2, 1/2)$ are

$$\lambda_{\pm} = -\frac{\alpha}{\beta} \pm iA. \tag{2.30}$$

As the stability of a fixed point of a continuous dynamical system is determined by whether the real part of the eigenvalues is positive, it is easy to see that $(1/2, 1/2)$ is always stable for $\alpha > 0$. Therefore, the replicator dynamics with finite memory always converges to the mixed strategy NE. When $\alpha = 0$ the fixed point becomes marginally stable, and the learning dynamics circles around the NE. This recovers a standard result in evolutionary game theory (Hofbauer and Sigmund, 1998). Note however that if $B \neq 0$ or $D \neq 0$ the position of the fixed point in the strategy space becomes dependent on α/β, and can be arbitrarily far from the mixed strategy NE.

2.A.2 Pure strategy fixed points stability

In order to study the properties of the pure strategy NE we need to consider the learning dynamics in the original coordinates (the pure strategies map into infinite elements in the transformed coordinates). The EWA dynamics reads (using (2.5) with $\delta = 1$, $\kappa = 1$ and the payoff matrix (2.1)):

$$\begin{aligned} x(t+1) &= \frac{x(t)^{1-\alpha} e^{\beta(ay(t) + b(1-y(t))}}{x(t)^{1-\alpha} e^{\beta(ay(t)+b(1-y(t))} + (1-x(t))^{1-\alpha} e^{\beta(cy(t)+d(1-y(t))}}, \\ y(t+1) &= \frac{y(t)^{1-\alpha} e^{\beta(ex(t) + f(1-x(t))}}{y(t)^{1-\alpha} e^{\beta(ex(t)+f(1-x(t))} + (1-y(t))^{1-\alpha} e^{\beta(gx(t)+h(1-x(t))}}. \end{aligned} \tag{2.31}$$

From Eq. (2.31) we can see that the pure strategies $(x, y) \in \{(0,0),(0,1),(1,0),(1,1)\}$ are all fixed points of the dynamics. Let us study their stability properties. We get a Jacobian

$$J = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}, \tag{2.32}$$

with

$$\begin{aligned} J_{11} &= \frac{(1-\alpha)(x-x^2)^\alpha e^{\beta(y(a-b-c+d)+b-d)}}{\left(x(1-x)^\alpha e^{\beta(y(a-b-c+d)+b-d)} - (x-1)x^\alpha\right)^2}, \\ J_{12} &= \frac{\beta(x-x^2)^{\alpha+1}(a-b-c+d)e^{\beta(y(a-b-c+d)+b-d)}}{\left(x(1-x)^\alpha e^{\beta(y(a-b-c+d)+b-d)} - (x-1)x^\alpha\right)^2}, \\ J_{21} &= \frac{\beta(y-y^2)^{\alpha+1}(e-f-g+h)e^{\beta(x(e-f-g+h)+f-h)}}{\left(y(1-y)^\alpha e^{\beta(x(e-f-g+h)+f-h)} - (y-1)y^\alpha\right)^2}, \\ J_{22} &= \frac{(1-\alpha)(y-y^2)^\alpha e^{\beta(x(e-f-g+h)+f-h)}}{\left(y(1-y)^\alpha e^{\beta(x(e-f-g+h)+f-h)} - (y-1)y^\alpha\right)^2}. \end{aligned} \tag{2.33}$$

As it can be seen by taking the appropriate limits in Eqs. (2.33), for all profiles of pure strategies the Jacobian has infinite elements along the main diagonal and null elements along the antidiagonal. This means that the profiles of pure strategies — and in particular the pure strategy NE — are "infinitely" unstable.

However, when $\alpha = 0$ the pure strategy NE become stable. Consider the profile of pure strategies in which both players choose action s_1. This corresponds to $x = y = 1$, and gives a Jacobian

$$J = \begin{pmatrix} e^{-\beta(a-c)} & 0 \\ 0 & e^{-\beta(e-g)} \end{pmatrix}. \tag{2.34}$$

The eigenvalues can be seen on the main diagonal, and are stable when $a > c$ and $e > g$. Under these conditions (s_1^R, s_1^C) is a pure strategy NE. The argument is similar for all other pure strategy profiles.

2.A.3 Mixed strategy fixed points stability

We first consider claim *(i)*. Since $B = 0$, there is always a fixed point $(\tilde{x}^\star, \tilde{y}^\star) = (0,0)$. This fixed point is stable if (from Eq. 2.11)

$$\frac{\beta}{\alpha}\,|A| \leq 1. \tag{2.35}$$

So, as long as $\tilde{x}^\star = 0$ is the unique fixed point, it is stable.

We then consider claim *(ii)*, and in particular the lower bound, $\frac{\beta}{\alpha}\,|A| \to 1^+$. Further to the central fixed point, there are two lateral fixed points $\tilde{x}^\star = \pm\epsilon$, where ϵ is an arbitrarily

small number. Thanks to the symmetry of the game, we focus on a profile of mixed strategies given by $(\tilde{x}^\star, \tilde{x}^\star)$. To second order, $\cosh \tilde{x}^\star \approx 1 + (\tilde{x}^\star)^2/2$. The stability condition becomes

$$\frac{\alpha}{\beta}\left(1+\frac{(\tilde{x}^\star)^2}{2}\right)\left(1+\frac{(\tilde{x}^\star)^2}{2}\right) - |A| \geq 0, \tag{2.36}$$

i.e.

$$(\tilde{x}^\star)^2 \geq \frac{\beta}{\alpha}|A| - 1. \tag{2.37}$$

Now, we Taylor expand $\Psi(\tilde{x}^\star)$ (defined in Section 2.3.2) to third order (first order would just yield $\tilde{x}^\star = 0$)) and solve $\tilde{x}^\star = \Psi(\tilde{x}^\star)$. Apart from the null solution, we get

$$(\tilde{x}^\star)^2 = \frac{3\left(\frac{\beta^2 A^2}{\alpha^2} - 1\right)}{\frac{\beta^2 A^2}{\alpha^2}\left(1+\frac{\beta^2 A^2}{\alpha^2}\right)}. \tag{2.38}$$

It is easily checked that for $\frac{\beta}{\alpha}|A| \to 1^+$, the condition (2.37) is satisfied: the fixed points whose components are the "lateral solutions" are stable. Therefore, there is a supercritical pitchfork bifurcation at the value $\frac{\beta}{\alpha}|A| = 1$.

The upped bound, namely $\frac{\beta}{\alpha}|A| \to \infty$, is easily dealt with. As discussed in Section 2.3.2, in this limit the fixed point $\tilde{x}^\star$ is given by $\tilde{x}^\star \approx \pm\frac{\beta}{\alpha}|A|$. Now, for $\frac{\beta}{\alpha}|A| \to +\infty$ the hyperbolic cosine can be approximated by

$$\cosh\left(\frac{\beta}{\alpha}|A|\right) \approx \exp\left(\frac{\beta}{\alpha}|A|\right)/2. \tag{2.39}$$

We can rewrite the stability condition as:

$$\frac{4\beta}{\alpha}|A|\exp\left(-2\frac{\beta}{\alpha}|A|\right) \leq 1. \tag{2.40}$$

For $\frac{\beta}{\alpha}|A| \to \infty$, the left hand side of the above equation goes to zero, so the inequality obviously holds.

Finally, the proof of *(iii)* is identical to the proof of the upper bound for $\frac{\beta}{\alpha}|A|$, in that the same arguments apply to sufficiently large values of $\frac{\beta}{\alpha}|B|$ (provided that $B \gg A$).

2.A.4 Fixed points of reinforcement learning

The fixed points are obtained setting $x(t+1) = x(t) = x^\star$ and $y(t+1) = y(t) = y^\star$ in Eq. (2.5), with $\alpha = 0$ and $\kappa = 1$. This gives

$$\begin{aligned} x^\star &= \frac{x^\star e^{\beta[\delta+(1-\delta)x^\star](ay^\star+b(1-y^\star))}}{x^\star e^{\beta[\delta+(1-\delta)x^\star](ay^\star+b(1-y^\star))} + (1-x^\star)e^{\beta[\delta+(1-\delta)(1-x^\star)](cy^\star+d(1-y^\star))}}, \\ y^\star &= \frac{y^\star e^{\beta[\delta+(1-\delta)y^\star](ex^\star+f(1-x^\star))}}{y^\star e^{\beta[\delta+(1-\delta)y^\star](ex^\star+f(1-x^\star))} + (1-y^\star)e^{\beta[\delta+(1-\delta)(1-y^\star)](gx^\star+h(1-x^\star))}}. \end{aligned} \tag{2.41}$$

It is easily checked that all pure strategy profiles are fixed points. Four additional solutions can be found noticing that when either $x^\star$ or $y^\star$ are 0 or 1, the respective equation holds as an identity. It is then possible to find the other solution by solving the other equation. These give fixed points

$$\begin{aligned}
(0, y_1) &= \left(0, \frac{(1-\delta)h + \delta(h-f)}{(1-\delta)(f+h)}\right), \\
(1, y_2) &= \left(1, \frac{(1-\delta)g + \delta(g-e)}{(1-\delta)(g+e)}\right), \\
(x_1, 0) &= \left(\frac{(1-\delta)d + \delta(d-b)}{(1-\delta)(d+b)}, 0\right), \\
(x_2, 1) &= \left(\frac{(1-\delta)c + \delta(c-a)}{(1-\delta)(c+a)}, 1\right).
\end{aligned} \tag{2.42}$$

Of course for these solutions to exist it has to be $0 < y_1, y_2, x_1, x_2 < 1$. The two final solutions (x_3, y_3) and (x_4, y_4) are obtained when the arguments of the exponentials in Eq. (2.41) are identical, i.e. when

$$\begin{aligned}
(\delta + (1-\delta)x^\star)(ay^\star + b(1-y^\star)) &= (\delta + (1-\delta)(1-x^\star))(cy^\star + d(1-y^\star)), \\
(\delta + (1-\delta)y^\star)(ex^\star + f(1-x^\star)) &= (\delta + (1-\delta)(1-y^\star))(gx^\star + h(1-x^\star)).
\end{aligned} \tag{2.43}$$

The expression for these solutions is very complicated and uninsightful. We only report this expression for $\delta = 0$ and symmetric games,

$$x^\star = y^\star = \frac{b - c + 2d \pm \sqrt{b^2 - 2bc + c^2 + 4ad}}{2(b - a + d - c)}. \tag{2.44}$$

We do not report the eigenvalues of other fixed points than the pure strategy profiles (Eq. 2.19) because their expression is complicated and uninsightful.

Chapter 3

Best reply structure and equilibrium convergence in generic games

3.1 Overview

In this chapter we introduce a theory that characterizes convergence of a family of learning algorithms in two-player games with any number of actions per player. Differently from Chapter 2, in which we restricted our attention to games with two actions per player, we cannot provide an exact characterization of the asymptotic behavior of these learning algorithms. We therefore calculate how typical convergence is over ensembles of two-player, normal-form, randomly-generated games, using an approximate measure based on the best reply structure of a game.

In Section 3.2 we explain why studying randomly-generated games and using approximate measures, both unconventional approaches in game theory, is a productive research agenda, by giving examples of how it has been useful in other fields. We then present some arguments giving some intuition about why the best reply structure of a game predicts non-convergence of some learning algorithms in Section 3.3. The key idea is that the relative "size" of best reply cycles vs. fixed points approximates the basin of attraction of unstable dynamics. This is remarkable given that the trajectories of learning algorithms do not closely match best reply cycles, and given that the functional form of learning algorithms is different from best reply dynamics. We perform extensive numerical simulations and give quantitative evidence for this correspondence in Section 3.4.

In Section 3.5 we study whether best reply cycles become prevalent as some properties of the games change, illustrating the effect of two different constraints that represent deviations from well-known classes of games. We show that best reply cycles are dominant in complicated and competitive games, indicating that in this case learning dynamics

are unlikely to reach equilibrium. In Section 3.6 we develop a combinatorial approach to analytically compute the frequency of best reply cycles in the case of uncorrelated payoffs.

Finally, in Section 3.7 we discuss perspectives for this work, some of which are already undergoing.

The results presented in this chapter have been published in Pangallo et al. (2019).

3.2 Why should we study randomly-generated games?

The standard approach to the problem of equilibrium convergence in game theory is to focus on *classes* of games with special mathematical properties, selected as stylized models of real-world scenarios. For example, in *potential games* (Monderer and Shapley, 1996b) all differences in payoffs for unilaterally changing strategy can be expressed using a global potential function; congestion games (Rosenthal, 1973) belong to this class, so potential games can be used as a stylized model of traffic. Most learning algorithms converge to a Nash or correlated equilibrium in potential games, as well as in dominance-solvable (Nachbar, 1990), coordination (Foster and Young, 1998), supermodular (Milgrom and Roberts, 1990) and "weakly acyclic" (Arieli and Young, 2016) games (see Section 1.2.3 for definitions). Studying classes of games with special properties is really useful in situations such as mechanism design, in which one can choose the game to be played. In these situations it may also be possible to design the learning algorithm that the players are going to use, and choose the one that most likely will converge to equilibrium (e.g. the ones in Section 1.3.7). Examples of situations in which this is useful include online auctions and internet routing (Easley and Kleinberg, 2010).

There are other problems, however, where the game and the learning algorithm are not designed, but rather are dictated by the intrinsic nature of the set up. For example, a financial market can be viewed as a game with many possible actions corresponding to the assets that traders can buy or sell. The outcome might converge to an equilibrium or it might endogenously fluctuate. If the system is not designed to encourage convergence, which of these two behaviors should we expect?

To address this issue, we follow an approach that has been extremely productive in theoretical ecology and developmental biology and is widespread in physics. An example of this approach is the seminal paper in theoretical ecology by May (1972). He studied an ensemble of randomly generated predator-prey interactions which he took as a *null model* of a generic ecosystem. His key result was that random ecosystems tend to become more unstable as they grow larger. Of course, as May was well-aware, real ecosystems are not random, rather they are shaped by evolutionary selection and other forces. Many large ecosystems have existed for long periods of time, suggesting they are in fact stable. This

indicated that real ecosystems are not typical members of the random ensemble he studied. This contradiction raised the important question of precisely how they are atypical, and how and why they evolved to become stable. Forty five years later, properly answering this question remains a subject of active research. For example, Johnson et al. (2014) recently found that real ecosystems have a property that they call *trophic coherence*, and showed that incorporating this property as a constraint on the ensemble of randomly generated ecosystems ensures stability.

Here we apply a similar approach to game theory, taking an ensemble of randomly generated two-player games as a null model. For reasons of tractability we study normal form games, taking advantage of the fact that it is possible to systematically enumerate all possible games. The null model is refined by adding constraints that can be varied to understand their effect on convergence to equilibrium. Here we study in detail a parameter Γ that tunes the correlation of the payoffs to the two players. This regulates the intensity of competition in a game, and encompasses zero-sum games as a special case for $\Gamma = -1$. We also sketch how one might construct other constraints, for example to study deviations from potential games. With this approach it is possible to see how deviations from particular classes of games affect the stability of the learning dynamics.

Randomly generated games and general learning algorithms do not have mathematical properties that allow exact solutions. Motivated by this, we develop a formalism to obtain the approximate probability of convergence as a function of a simple indicator. An analogy to fluid dynamics clarifies how such approximate solutions can be useful. As a fluid is driven away from equilibrium, it typically makes a transition from stable (or *laminar*) flow to unstable (or *turbulent*) flow. There is no complete theoretical understanding of the solutions for the Navier-Stokes equations that describe fluid dynamics, but this transition can nonetheless be crudely characterized in terms of a nondimensional parameter called the Reynolds number (Morriss and Evans, 2013). Larger Reynolds number means higher likelihood of turbulence. Although this is imprecise – it is just a rule of thumb – it is nonetheless very useful. Our analogous estimate for games does not have a simple closed form, but it has similar predictive power. Another analogy to our approach is the theory of qualitative stability in theoretical ecology (May, 1973). Many models in ecology consider the magnitude of the interactions between different species in a food web. For example, such models consider how much grass is eaten by rabbits, and how many rabbits are eaten by foxes. The qualitative stability approach instead considers only the sign of the predator-prey relations – that rabbits eat grass and foxes eat rabbits. This makes it possible to obtain an approximate assessment of the stability of an ecosystem just from the topological

properties of its food web. As with the theory of qualitative stability, our approach only depends on the topological properties of the game and not on the details of the payoffs.

3.3 How the best reply structure shapes learning dynamics

Table 3.1: Terminology

Best reply	Action that gives the best payoff in response to a given action by an opponent
Best reply structure	Arrangement of the best replies in the payoff matrix
Best reply matrix	Derived payoff matrix, with one for the best reply to each possible action of the opponent and zero everywhere else
Best reply dynamics	Simple learning algorithm in which the players myopically choose the best reply to the last action of their opponent
Best reply k-cycle	Closed loop of best replies of length k (each player moves k times)
Best reply fixed point	Pure Nash Eq., i.e. the action for each player that is a best reply to the action of the other player
Best reply vector $\boldsymbol{v}$	List of the number of distinct attractors of the best reply dynamics, ordered from longest cycles to fixed points
Free action / free best reply	Best reply to an action that is neither part of a cycle or fixed point
Best reply configuration	Unique set of best replies by both players to all actions of their opponent

The simple indicator that we develop to obtain an approximate probability of convergence is based on best reply dynamics. Some definitions on best reply dynamics have been given in Section 1.3.1; in Table 3.1 we recall these definitions, together with others that will be added in Section 3.6.

Our working hypothesis is that the best reply structure of a game influences the convergence properties of a wide variety of learning algorithms that have empirical support from experiments with human players. We consider six learning algorithms that span different information conditions and levels of rationality, defined in Sections 1.3.2 through 1.3.6. These are reinforcement learning, fictitious play, replicator dynamics, experience-weighted attraction (with and without noise) and level-k learning. We specifically do *not* study algorithms that are designed to have good convergence properties, which may be very different from the ones we study here and for example converge to correlated equilibria independently of the best reply structure.

More specifically, our hypothesis is that the presence or absence of best reply cycles is correlated with the stability of the learning dynamics, even if the learning trajectories may not follow best reply cycles in detail. Learning is more likely to converge when there are no best reply cycles and less likely to converge when they are present. Note that *we are*

not assuming that players follow best reply dynamics. Rather, we hypothesize that the best reply structure of the payoff matrix constitutes a first-order skeleton, forming the backbone of the game the players are trying to learn, which is useful for understanding the convergence of many learning algorithms.

We cannot prove this analytically in generic games, but it is supported by anecdotal examples. It is impossible to present these in detail but in this section we present a few representative samples to motivate the correspondence. In the next section we present more quantitative evidence.

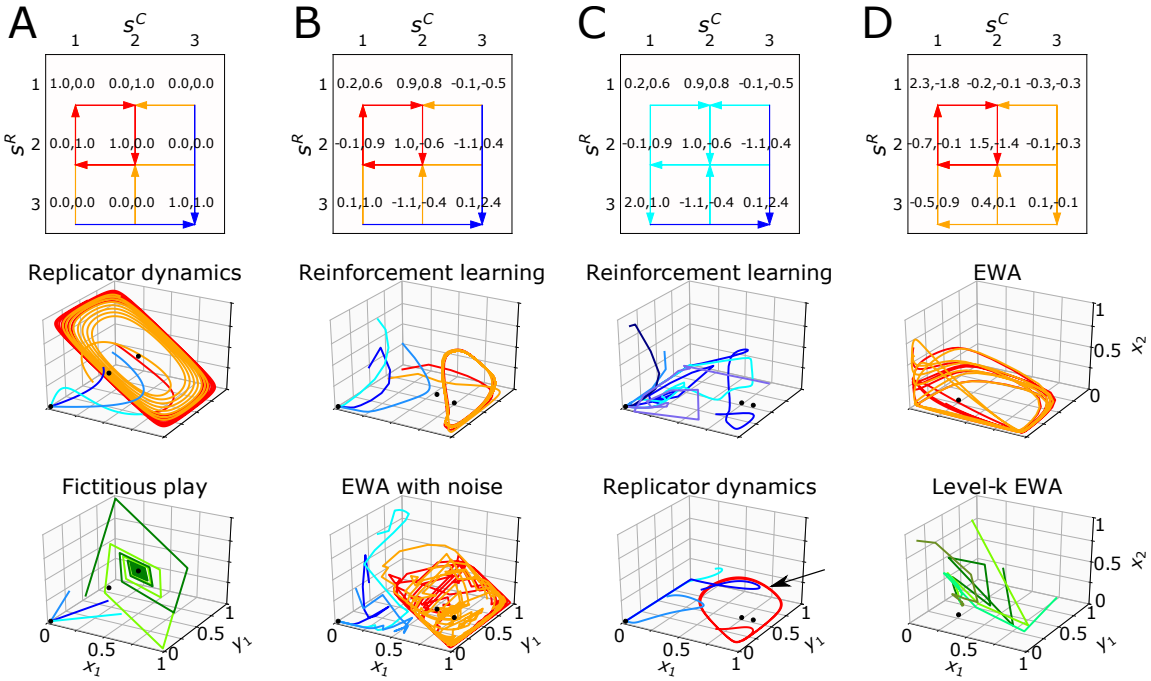


Figure 3.1: **How the best reply structure influences convergence of six learning algorithms.** For each payoff matrix A–D, we show the dynamics of two learning algorithms. We plot the probabilities x_1 and x_2 for Row to play actions $s^R = 1$ and $s^R = 2$ respectively, and the probability y_1 for Column to play $s^C = 1$. The black dots are the positions of the Nash equilibria: for payoff matrices A–C the dot at $x_1 = x_2 = y_1 = 0$ is the pure strategy Nash equilibrium; all other dots are mixed strategy equilibria. The trajectories are colored in shades of blue if they converge to a pure equilibrium; in shades of green if they converge to a mixed equilibrium; and in shades of red if they do not converge.

To develop intuition into whether and why the best reply structure could predict convergence of the learning algorithms above, in Figure 3.1 we analyze four games with $N = 3$, showing some learning trajectories in a 3-dimensional projection of the 4-dimensional strategy space (there are six components of the mixed strategy vectors with two normalization constraints). The axis labels x_1 and x_2 are the probabilities of player Row to play actions $s^R = 1$ and $s^R = 2$ respectively, and y_1 is the probability for Column to play $s^C = 1$. The corners of the strategy space $(x_1, y_1, x_2) = (1, 1, 0)$ and $(0, 1, 1)$ correspond to the action profiles $(1, 1)$ and $(2, 1)$ respectively.

In panel A we consider a best reply matrix with a 2-cycle and a single pure strategy Nash equilibrium. Our first illustration uses replicator dynamics. The attractors of the learning algorithm closely mimic the attractors of the best reply dynamics; all trajectories go to either the fixed point or the limit cycle, depending on the initial condition. The limit cycle in this example corresponds to the best reply cycle as we always have $y_3 = 0$. (Coordinate y_3 is not shown due to the 3-dimensional projection of the 6-dimensional system.) Reinforcement learning, EWA and EWA with noise behave similarly. In contrast, fictitious play always converges to a fixed point, either to the pure strategy Nash equilibrium or to the mixed strategy equilibrium in the support of the cycle, depending on initial conditions. We never observe it converging to the other mixed NE, which corresponds to the combination of the two attractors. In Section 3.4.2, we show that in generic games with $N = 10$ fictitious play is more likely to converge to mixed equilibria in the support of a best reply cycle than to other mixed equilibria, with respect to the proportions of existing mixed equilibria. Level-k EWA also converges (close) to the same mixed equilibrium. As we will show quantitatively in the following section, level-k EWA behaves like fictitious play in games with few actions, and more like the other four algorithms in games with many actions. For best reply matrices it is not surprising that the learning dynamics mimic the best reply structure, but because the learning algorithms have free parameters and memory of the past it is not obvious that they should do so this closely.

The payoff matrix in panel B has the same best reply structure as in A, but has generic payoffs. To show a wider variety of learning algorithms, we illustrate the learning dynamics this time with reinforcement learning and EWA with noise. In both cases the learning trajectories either converge to the pure strategy Nash equilibrium, or do not converge. Reinforcement learning converges to a limit cycle that is related to the best reply cycle, even if the trajectory is distorted and penetrates more deeply into the center of the strategy space, and there is a similar correspondence for EWA with noise. As in panel A, replicator dynamics and EWA behave similarly, while fictitious play and level-k EWA converge to or close to mixed equilibria.

To see what happens once the best reply cycle is removed, in panel C we consider the same payoff matrix in B, except that the payoff for Row at $(3, 1)$ is 2.0 instead of 0.1. In this case the only attractor of the best reply dynamics is the pure strategy NE at $(3, 3)$. For all initial conditions reinforcement learning converges to the pure strategy Nash equilibrium, illustrating how removing the best reply cycle makes the fixed point globally stable. However, for replicator dynamics there exist some initial conditions leading instead to a limit cycle, suggesting that a small basin of attraction for unstable learning dynamics may still exist even in the absence of best reply cycles. (Fictitious play and level-k EWA

behave like reinforcement learning, while EWA and EWA with noise behave like replicator dynamics.)

This case illustrates an important example of how the learning dynamics can qualitatively deviate from the best reply dynamics. When the replicator dynamics reaches the point $(x_1, y_1, x_2) = (0.86, 0.87, 0.14)$, indicated by a black arrow in the figure, the expected payoff for Row is 0.25. If Row switches to $s^R = 3$ (his best reply), his payoff raises to 1.60. If instead Row switches to $s^R = 1$ Row's expected payoff still increases to 0.29, so we say that action 1 is a *better reply*. With replicator dynamics the probability for actions corresponding to better replies grows over time, although at a lower rate than best replies. Here because of past play x_3 is very small (of the order of 10^{-53}), so x_1 increases to 1 before x_3 can reach a similar value, and the dynamics gets trapped in the cycle shown in the figure.

In panel D we consider a payoff matrix with a 2-cycle and no pure strategy Nash equilibria. EWA reaches a chaotic attractor for all initial conditions. Note that this attractor shares some of the structure of the best reply cycle, visiting the two corners of the strategy space corresponding to $(x_1, y_1, x_2) = (1, 0, 0)$ and $(1, 1, 0)$ closely and the other two more loosely. Level-k EWA, in contrast, converges close to the mixed Nash equilibrium in the center of the chaotic attractor.[29] The other algorithms behave similarly: both EWA with noise and replicator dynamics reach a chaotic attractor, while reinforcement learning follows a limit cycle. Fictitious play fluctuates close to the mixed equilibrium, with each component of the mixed strategy vector at a maximum distance of 0.03 from the equilibrium point.

We thus see that, despite the fact that the learning dynamics of these algorithms do not mimic the best reply dynamics in detail, in most cases remnants of it are seen in the learning dynamics. Even if the correspondence is not exact, the best reply structure influences the learning dynamics.

3.4 Quantitative evidence on the effect of the best reply structure

We now quantitatively test the hypothesis that the presence of best reply cycles is positively correlated to nonconvergence, at least for these learning algorithms. To do this we first prove a very useful theorem relating the configuration of cycles and fixed points to the relative size of the basin of attraction for unstable best reply dynamics. Recalling that n_k is the number of best reply cycles of length k, let $C = \sum_{k=2}^{N} n_k k$ be the number of actions that are part of best reply cycles. The quantity

$$\mathcal{F}(\boldsymbol{v}) = C/(C + n_1), \tag{3.1}$$

[29] Due to finite memory and finite payoff sensitivity, EWA, EWA with noise and level-k EWA cannot reach mixed equilibria exactly. This is shown in Chapter 2.

measures the relative number of actions on attractors that are part of cycles vs. fixed points. For the example of Figure 1.1, where there is one fixed point and one cycle, $\mathcal{F}(0,0,1,1) = 2/3$.

In fact this also measures the relative size of the basins of attraction of cycles vs. fixed points under best reply dynamics. This is not true for a given individual game, but it is true for the average over the ensemble of all games with a given best reply vector $\boldsymbol{v}$. For example for the game shown in Figure 1.1, 5/8 of the initial conditions lead to the cycle and 3/8 lead to the fixed point. However if one averages over all possible 4×4 games that have one best reply cycle of length 2 and a single pure strategy Nash equilibrium and no other best reply attractors (giving an equal weighting to each possible game), the relative size of initial conditions leading to the best reply cycle is exactly 2/3.

Examples and a proof are given in Pangallo et al. (2019).

3.4.1 Main results

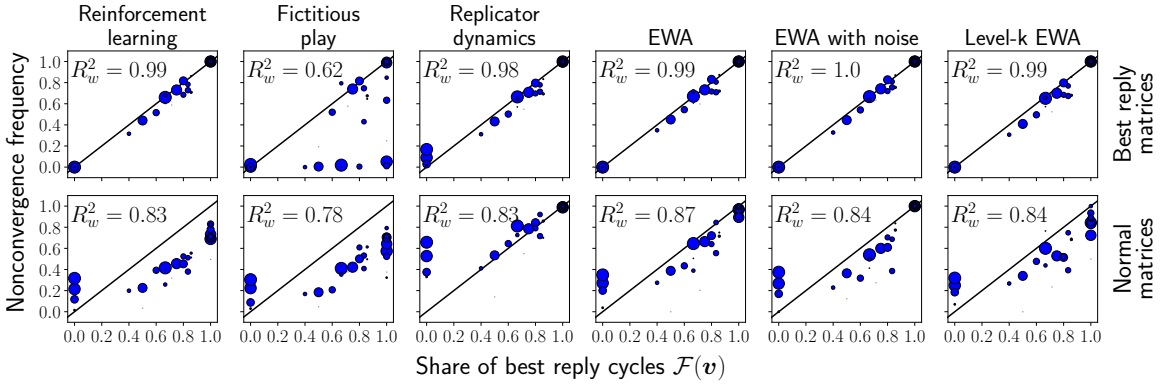


Figure 3.2: **Test for how well the best reply structure predicts non-convergence for six learning algorithms.** We generate 1000 random payoff matrices with $N = 20$ actions and for each of these we simulate learning 100 times based on random initial conditions. The same process is repeated using the best reply matrix associated with each of the 1000 random games. Each circle corresponds to a specific best reply vector $\boldsymbol{v}$. Its size is the logarithm of the number of times a payoff matrix with $\boldsymbol{v}$ was sampled. The horizontal axis is the share of best reply cycles $\mathcal{F}(\boldsymbol{v})$. For example, the largest circle at $\mathcal{F}(\boldsymbol{v}) = 0.66$ corresponds to $\boldsymbol{v} = (0, \ldots, 0, 1, 1)$. The vertical axis gives the frequency of non-convergence in the simulations, averaged over all payoff matrices and initial conditions having the same $\boldsymbol{v}$. The top row shows results for the best reply matrices and the bottom row shows results using normally distributed payoffs. The identity line is plotted for reference.

To test the relationship between best reply dynamics and the convergence properties of our family of learning algorithms we generate games at random. This is done by populating each payoff bi-matrix with $2N^2$ normally distributed numbers; as discussed in the Appendix, Section 3.A.7 the normal distribution is the maximum entropy distribution and so is the natural choice. Here we let the payoffs of the two players be uncorrelated. We then simulate the learning process of the players in a repeated game, holding the payoff

matrix fixed. Convergence to pure and mixed strategy Nash equilibria is checked using the criteria explained in Appendix 3.A. We generate 1000 different games; for each game we simulate all six learning algorithms starting from 100 different initial conditions. As a point of comparison, we repeat the entire process using the best reply matrices associated with each of the 1000 randomly generated games. Results with $N = 20$ are reported in Figure 3.2 and results with $N = 5$ and $N = 50$ are given in Figures 3.4 and 3.5, later in this section.

Figure 3.2 compares the share of best reply cycles $\mathcal{F}(\boldsymbol{v})$ to the nonconvergence frequency. In order to test the relationship between the best reply vector $\boldsymbol{v}$ and the learning dynamics, we group together the results for payoff matrices with the same $\boldsymbol{v}$ and plot a circle whose radius is proportional to the logarithm of the number of times this was sampled. We place each best reply vector on the horizontal axis according to its share of best reply cycles $\mathcal{F}(\boldsymbol{v})$. On the vertical axis we plot the average frequency of non-convergence for results with this best reply vector. Thus if the best reply structure perfectly predicts the rate of convergence of the other learning algorithms, all circles should be centered on the identity line. We estimate the weighted correlation coefficient R^2_w using weights corresponding to the number of times each best reply vector was sampled.

The results when we simulate using the best reply matrices are shown in the top row of Figure 3.2. The correlation is nearly one for all the algorithms except fictitious play. On one hand, given that best reply matrices do not have better replies, this may not seem too surprising. But on the other hand, these learning algorithms all use memory of the past (beyond the most recent action) and most of them have free parameters, while best reply dynamics has neither of these. Given that, it is remarkable that the size of the basins of attraction of learning algorithms with different functional forms are so well approximated by best reply dynamics. The failure for fictitious play is due to its strong tendency to converge to mixed strategy NE, though even in this case it typically fails to converge for longer cycles.[30] The results in the top panels of Figure 3.2 reinforce the intuition that for best reply matrices the attractors of the learning algorithms closely match the best reply dynamics for all algorithms except fictitious play.

The case with general payoffs generated from a Normal distribution is shown in the bottom row of Figure 3.2. The correlations are still very strong, with weighted correlation coefficient $R^2_w = 0.78$ for fictitious play and $R^2_w \approx 0.84$ for all of the other algorithms. The best reply dynamics are not a perfect predictor of convergence but the predictions are still good enough to be very useful.

[30] Fictitious play converges in most cases in the presence of 2-cycles, but fails to converge with 3-cycles, in line with Miyazawa (1961) and Shapley (1964). See Section 1.3.3.

This analysis also gives clues about some of the other factors that cause non-convergence. For example, even when $\mathcal{F}(\boldsymbol{v}) = 0$, indicating that best reply cycles are absent, convergence is not certain. This is evident from the vertical column of circles on the left of each figure. This column corresponds to best reply vectors with no cycles, i.e. those of the form $\boldsymbol{v} = (0, \ldots, 0, 0, x)$, where $x = 1, 2, 3, 4$ is the number of distinct fixed points. The highest circle corresponds to a single fixed point, the one below it to two fixed points, etc. In the case where there is a unique pure Nash equilibrium and no cycles, the nonconvergence frequency is typically around 35%, dropping to about 20% if there are two pure equilibria. The presence of multiple pure equilibria makes the existence of better reply cycles like the one in Figure 3.1C less likely. Conversely, there are also cases of convergence when there are cycles without any fixed points. This corresponds to the column of vertical circles on the right, with $\mathcal{F}(\boldsymbol{v}) = 1$, for level-k EWA, fictitious play and reinforcement learning.

Non-convergence in the presence of a unique pure Nash equilibrium with no cycles is particularly pronounced for replicator dynamics, where nonconvergence happens about 60% of the time when there is a single fixed point. As we explain in Appendix 3.A, this is in part due to numerical limitations that become more serious as N grows, so we will drop observations of replicator dynamics if $N \geq 50$. Convergence in the absence of fixed points is entirely due to convergence to mixed strategy Nash equilibria. This is discussed in Section 3.4.2.

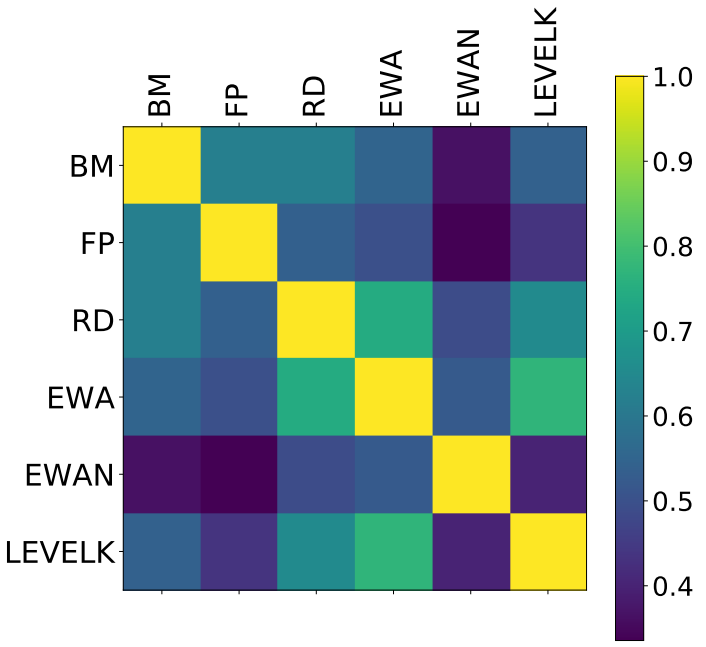


Figure 3.3: **Correlation matrix of the co-occurrence of non-convergence for the six learning algorithms**: Bush-Mosteller (BM) learning, Fictitious Play (FP), Replicator Dynamics (RD), Experience-Weighted Attraction (EWA), EWA with Noise (EWAN) and level-k learning (LEVELK).

In Figure 3.3 we show the correlation matrix of the convergence of the six learning algorithms. We find that the three most correlated algorithms are replicator dynamics, Experience-Weighted Attraction (EWA) and level-k learning. The two least correlated algorithms are fictitious play and EWA with noise. The correlation ranges between 0.35 and 0.85, suggesting a relatively strong heterogeneity among the six algorithms, consistently with the examples shown in Figure 3.1.

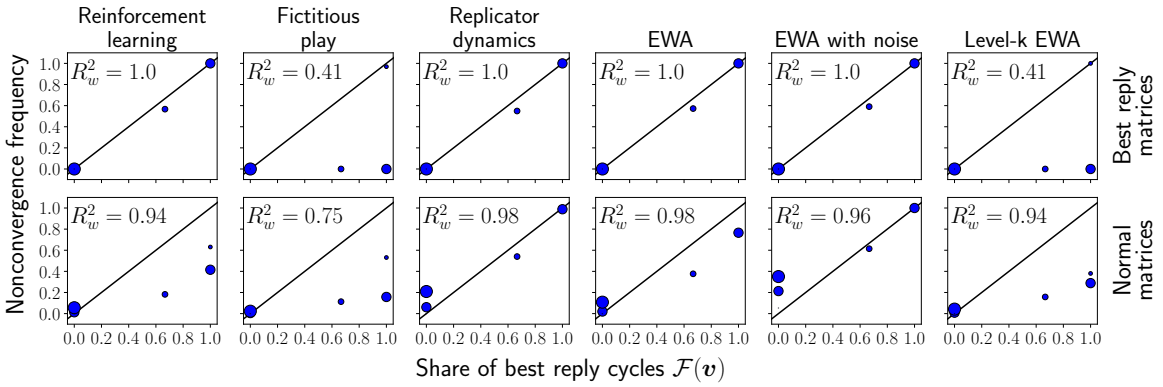


Figure 3.4: **Test for how well the best reply structure predicts non-convergence with** $N = 5$, instead of $N = 20$ as in Figure 3.2.

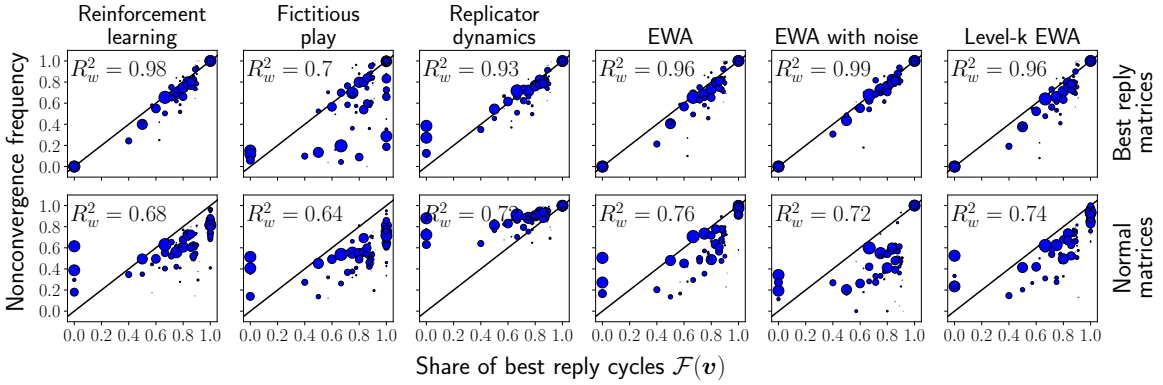


Figure 3.5: **Test for how well the best reply structure predicts non-convergence with** $N = 50$, instead of $N = 20$ as in Figure 3.2.

All results up to this point apply to $N = 20$. Are results robust to other values of N? As can be seen in Figures 3.4 and 3.5 for $N = 5$ and $N = 50$ respectively, the overall pattern is similar, but there are some differences.

For $N = 5$ the correlation is stronger than with $N = 20$, and the values of the weighted correlation coefficient are even larger than 0.9 in normal payoff matrices. We conjecture that this is due to a higher share of the actions that are part of cycles and fixed points. Indeed, for $N = 5$ the most common best reply vector with cycles is $\boldsymbol{v} = (0, 0, 0, 1, 0)$, so the actions that are part of the cycle are 2/5. On the other hand, in a best reply vector

with a 2-cycle and $N = 20$ the actions that are part of the cycle are $2/20$, so the payoffs that are not best replies have more importance and the non-perfect correspondence due to better replies is more severe.

An interesting detail is that level-k learning converges in most cases. Inspection of individual simulation runs suggests that by anticipating the actions of their opponent, the players are less likely to get stuck in periodic cycles and converge instead to mixed strategy equilibria. This finding is consistent with the examples shown in Figure 3.1, in which level-k learning was always converging close to mixed equilibria. It is also consistent with the panel with $N = 3$ in Figure 3.8, to be shown in the next section. This raises the interesting question of why having forward-looking players enhances convergence when N is small but does not do so when N is large.

For $N = 50$ we observe the opposite pattern than with $N = 5$: the correlation becomes weaker (but is still larger than 0.6 in all cases). This effect is most likely caused by a smaller share of actions that are part of cycles or fixed points (the most common best reply vector is $\boldsymbol{v} = (0, \dots 0, 1, 1)$, involving only $3/50$ of the actions). Better replies probably play a more important role. However, we cannot exclude measurement error, especially in the case of replicator dynamics, as explained in Appendix 3.A. Indeed, with $N = 50$ replicator dynamics converges only in a minority of cases, even in the absence of best reply cycles, suggesting that the results are driven by numerical limitations. For this reason, we exclude results with $N \geq 50$ in the case of replicator dynamics.

In summary, even if $\mathcal{F}(\boldsymbol{v})$ ignores better replies and underestimates the convergence to mixed strategy Nash equilibria, there exists a robust correlation between the average probability of convergence and the share of best reply cycles. This indicates that the best reply structure and the stability of these algorithms are closely linked.

3.4.2 Mixed equilibria and the best reply structure

An issue that has been sidestepped up to this point but that is nonetheless very important is the role played by mixed strategy Nash equilibria. In this section we discuss the relation between mixed equilibria and the best reply structure. In particular, we explain why mixed equilibria reduce the predictive power of the best reply structure on the convergence frequency of learning algorithms, yet this is true to a small extent so that the predictions from our theory are still useful.

Once the attractors of the best reply dynamics are known it is trivial to list all the mixed strategy Nash equilibria of the best reply matrix. There is one mixed NE associated with each best reply cycle and there is one associated with all possible combinations of the cycles and fixed points. The mixed strategy NE associated with a given cycle corresponds

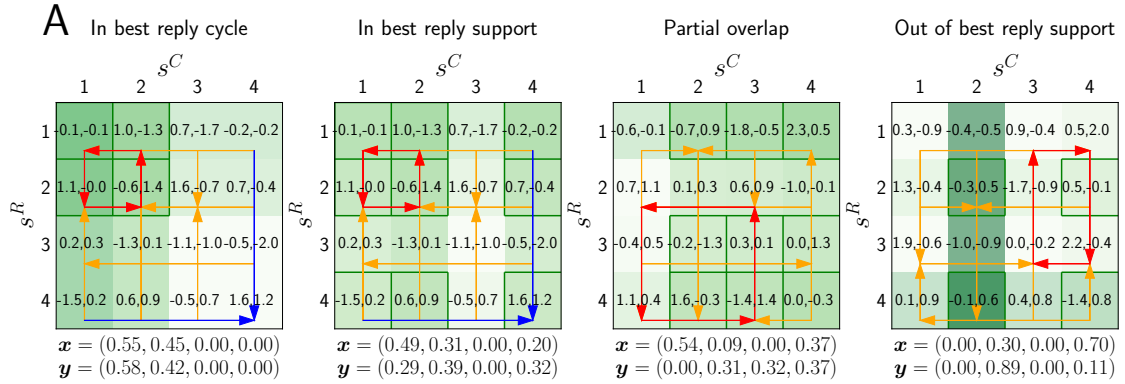
to randomly playing each action with the frequency with which it is visited on the cycle. For example, the mixed strategy equilibrium associated with the best reply cycle in Figure 1.1B is $\boldsymbol{x}, \boldsymbol{y} = (0.5, 0.5, 0, 0), (0.5, 0.5, 0, 0)$. The mixed strategy NE associated with each possible combination of cycles and fixed points corresponds to playing the average over the combined action sets. For the example in Figure 1.1B the mixed NE associated with the combination of the cycle and the fixed point is $\boldsymbol{x}, \boldsymbol{y} = (0.33, 0.33, 0.33, 0), (0.33, 0.33, 0.33, 0)$. For the best reply matrix there are no other mixed strategy NE. These statements are made more precise and proved in Pangallo et al. (2019). (We do not report the proof here because it is mainly the work of Torsten Heinrich, coauthor of the paper.)

In moving from the best reply matrix to the original game with payoffs generated from a Normal distribution, some of the mixed NE may survive and others may not, and new mixed NE may be introduced. For example in moving from the best reply matrix in Figure 1.1A to the payoff matrix in Figure 1.1A, none of them survive, and there are two mixed equilibria at $\boldsymbol{x}, \boldsymbol{y} = (0.32, 0, 0, 0.68), (0.36, 0.64, 0, 0)$ and at $\boldsymbol{x}, \boldsymbol{y} = (0.26, 0.15, 0, 0.59), (0.32, 0.24, 0.44, 0)$, which have no relation to those of the associated best reply matrix.

To understand the extent to which the properties of mixed equilibria that are valid in best reply payoff matrices hold in normal payoff matrices, we randomly generate 1000 payoff matrices with $N = 10$, sampling all payoffs independently from a standard Gaussian distribution. By using the vertex enumeration method (Nisan et al., 2007) as implemented in the Python package `Nashpy` we find *all* mixed strategy Nash equilibria in all 1000 games.

In our sample, only 32% of mixed equilibria in the support of 2-cycles and 10% of mixed equilibria in the support of 3-cycles "survive" to the introduction of non-Boolean payoffs. There exist no mixed equilibria in the support of cycles of length greater than three.

Four possible classes of mixed equilibria are illustrated in the four panels of Figure 3.6A. The cells in the payoff matrices that are surrounded by a green square are the action profiles that are played with positive probability at the mixed equilibrium. In other words, they constitute the support of the mixed equilibrium. In the first case the support of the mixed equilibrium coincides with the support of the best reply cycle. In the second case it coincides with the combined support of the cycle and the pure strategy equilibrium. These first two cases also occur in best reply matrices, although in normal matrices it is no longer true that all actions are played with the same probability. In the third case only the action profile $(s^R, s^C) = (4, 3)$ – part of the support of the best reply cycle – overlaps with the support of the mixed equilibrium. In the fourth case, there is no overlap between the supports. The last two cases show that in normal payoff matrices there can be mixed equilibria that



B

Existence:	2.7%	3.7%	67.7%	25.9%
Convergence BM:	27.5%	1.0%	41.1%	30.4%
Convergence FP:	29.7%	1.3%	41.8%	27.2%

Figure 3.6: **Mixed strategy Nash equilibria classified in relation to the best reply structure.** (A) Examples of mixed equilibria $(\boldsymbol{x}, \boldsymbol{y})$ in different games. Each row of each payoff matrix is colored in shades of green whose intensity is proportional to the respective component of the mixed strategy vector $\boldsymbol{x}$ reported below. Darker shades of green indicate that action i is played with higher probability, white indicates that it is played with null probability. The same coloring scheme applies to columns, except that the intensity of green is proportional to $\boldsymbol{y}$. The action profiles (s_i^R, s_j^C) surrounded by green squares are the Cartesian product of the actions of Row and Column that are played with positive probability. (Due to the coloring scheme, other cells of the payoff matrix are colored in shades of green, but they are never played.) Mixed equilibria can be (i) in the support of a best reply cycle; (ii) in the combined support of best reply attractors; (iii) partially overlapping with the support of best reply attractors; (iv) out of the support of best reply attractors. (B) Statistics from a sample of 1000 randomly generated games with $N = 10$. For each of the four categories above we report: (i) the percentage of mixed equilibria that fall in each category, over all mixed equilibria in all 1000 games; (ii) the percentage of converging simulation runs of Bush-Mosteller (BM) reinforcement learning that converge to mixed equilibria in that category, over all simulation runs that converge to any mixed Nash equilibrium; (iii) same as (ii), but for fictitious play (FP).

have only something or nothing to do with the attractors of best reply dynamics. These equilibria depend on the details of the payoffs.

How common are mixed equilibria in the support of best reply attractors vs. partially overlapping/out of the support? In Figure 3.6B we show the percentage of existing mixed equilibria that belong to each category, over all existing mixed equilibria in 1000 payoff matrices with $N = 10$. Only 6.4% of the mixed equilibria are in the support of the best reply attractors, while 67.7% are partially overlapping and 25.9% are out of the support. This indicates that the vast majority of mixed equilibria is at best partially related to the best reply attractors.[31]

How much do these equilibria matter? For each of the 1000 payoff matrices we simulate Bush-Mosteller reinforcement learning and fictitious play (these are the algorithms that converge to mixed equilibria most frequently for $N = 10$) and record the category of the mixed

[31]The quantitative results are likely to depend on N. An exhaustive characterization of the relation between mixed equilibria and the best reply structure as a function of N is beyond the scope of this paper. We choose $N = 4$ in Figure 3.6A for illustrative purposes only.

equilibria to which they converge. Interestingly, mixed equilibria in the support of best reply cycles are disproportionately more likely to be reached by these learning algorithms, with respect to the existence percentage. Indeed, Bush-Mosteller learning converges to such mixed equilibria 27.5% of the times, and fictitious play 29.7%, even if mixed equilibria in the support of best reply cycles are only 2.7% of the total. Convergence to other mixed equilibria in the support of best reply attractors is instead very unlikely, 1.0% and 1.3% for Bush-Mosteller learning and fictitious play respectively. These results suggest that in normal payoff matrices best reply cycles matter even when learning algorithms converge to mixed equilibria.

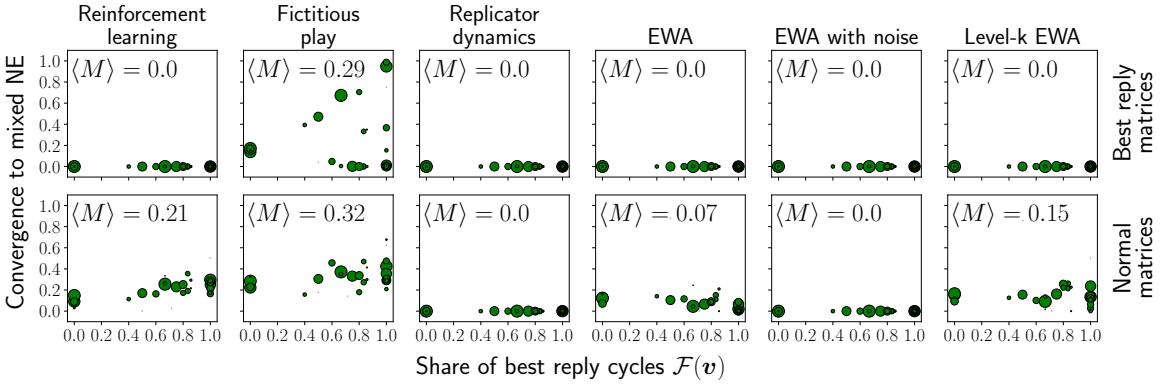


Figure 3.7: **Frequency of convergence to mixed strategy Nash equilibria** for $N = 20$. This figure should be interpreted as Figures 3.2, 3.4 or 3.5. $\langle M \rangle$ is the frequency of convergence to mixed equilibria averaged over all simulation runs.

To understand whether best reply cycles make it more or less likely that learning algorithms converge to mixed strategy Nash equilibria, in Figure 3.7 we show the frequency of convergence to mixed equilibria in the case of $N = 20$, using the same data as for Figure 3.3.[32] There are several interesting things to note:

1. The convergence to mixed equilibria is mostly independent of the best reply vector, i.e. it occurs independently of the presence of best reply cycles or fixed points.

2. In normal payoff matrices, convergence is highest in the case of fictitious play (32%), followed by Bush-Mosteller reinforcement learning and level-k EWA (21% and 15% respectively). Convergence to mixed equilibria is quite rare with EWA (7%) and impossible with two-population replicator dynamics (Gintis, 2000) and EWA with noise. In the latter case, this is because noise is very strong away from the faces of the strategy space, and so only pure strategy Nash equilibria can be stable (see Figure 3.19).

[32] We consider EWA, EWA with noise and level-k EWA convergent to a mixed equilibrium provided they reach a fixed point "sufficiently close" to the equilibrium, as detailed in Appendix 3.A.

3. In best reply payoff matrices, convergence to mixed equilibria never occurs for any learning algorithm except fictitious play (with $N = 20$). For fictitious play, convergence occurs in 95% of the cases for games with best reply vector $\boldsymbol{v}(\Pi) = (0, \ldots, 0, 1, 0)$, but only in 1% of the simulations in games with $\boldsymbol{v}(\Pi) = (0, \ldots, 1, 0, 0)$. (These correspond to the two largest circles at $\mathcal{F}(\boldsymbol{v}) = 1$ in the top panel of fictitious play.) These results are consistent with Miyazawa (1961), who proved convergence in all 2×2 games (featuring a 2-cycle); and with Shapley (1964), who proved non-convergence in a 3×3 game featuring a 3-cycle.

In sum, convergence to mixed Nash equilibria is almost independent of $\mathcal{F}(\boldsymbol{v})$, and so causes a constant offset. This is the reason that for $\mathcal{F}(\boldsymbol{v}) > 0$ the circles are below the identity line in Figure 3.2. Once these effects are taken into account, there is a strong proportionality between $\mathcal{F}(\boldsymbol{v})$ and the nonconvergence rate for all of the learning algorithms, including fictitious play.

3.5 Variation of the best reply structure

We now investigate the prevalence of best reply cycles and fixed points as we vary the properties of the games, and test the prediction of nonconvergence as we vary parameters. The parameters that we initially investigate are the number of possible actions N and the correlation Γ between the payoffs of the two players. Later we also consider a parameter ξ to interpolate between random and potential games.

As before we generate games at random, but we now impose constraints on the ensemble from which the games are drawn. To understand how Γ affects convergence, we generate payoff matrices by drawing from a bivariate normal distribution, so that the expected value of the product of the payoffs to players Row and Column for any given combination of actions is equal to Γ. A negative correlation, $\Gamma < 0$, implies that the game is competitive, because what is good for one player is likely to be bad for the other. The extreme case is where $\Gamma = -1$, meaning the game is zero-sum. In contrast $\Gamma > 0$ encourages cooperation, in the sense that payoffs tend to be either good for both players or bad for both players.

In Figure 3.8 we show how the share of best reply cycles varies with N and Γ. For a given value of N and Γ we randomly generate payoff matrices and compute the average share of best reply cycles $\langle \mathcal{F}(\boldsymbol{v}) \rangle_{N,\Gamma}$. We compare $\langle \mathcal{F}(\boldsymbol{v}) \rangle_{N,\Gamma}$ to the average frequency of non-convergence of the six learning algorithms. The agreement is generally good. The striking exception is fictitious play, which has a high convergence rate when N is large and Γ is negative. This is at odds with the other algorithms, which rarely converge in this

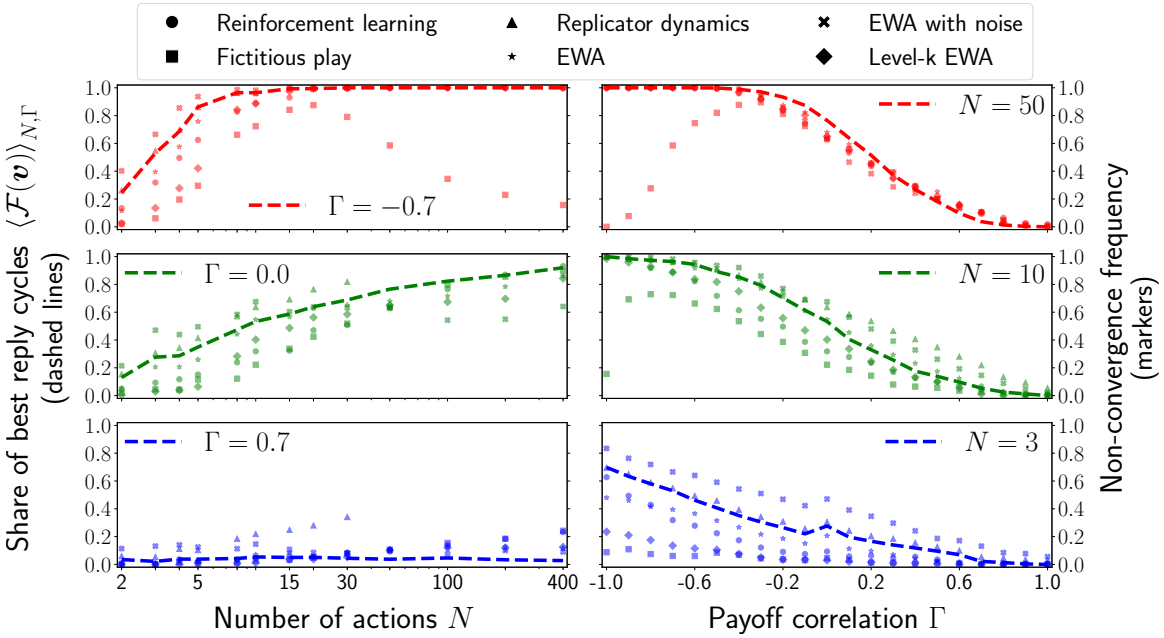


Figure 3.8: **How the share of best reply cycles predicts convergence as a function of the number of actions N and the competition parameter Γ.** Dashed lines are the average share of best reply cycles $\langle \mathcal{F}(\boldsymbol{v}) \rangle_{N,\Gamma}$ for those values of N and Γ. Markers are the fraction of simulation runs in which the learning algorithms do not converge.

parameter range, and agree closely with the prediction. Thus, as expected, fictitious play must be regarded somewhat differently than the other algorithms.

We would like to emphasize that these predictions are made *a priori* – they do not involve fitting any parameters. With the exception of fictitious play, the predictions fit the overall trends in the data, both estimating the magnitude correctly and capturing the functional dependence. For the other five algorithms the predictions are particularly good when N is large and Γ is strongly negative, with a near perfect fit for $N > 10$.

What are the conclusions about the circumstances in which convergence is likely as we vary N and Γ? When Γ is positive (meaning that the game is not competitive), convergence is almost guaranteed, regardless of N. But when Γ is strongly negative convergence is likely only when N is very small. At $\Gamma = -0.7$, for $N = 4$ the nonconvergence rate is roughly 70%, quickly rising to approach 100% by $N = 8$. Our analysis indicates that this is because complicated competitive games are dominated by cycles. In this region of the parameter space acyclic games are extremely rare. Therefore, dominance-solvable, coordination, potential and supermodular games are atypical.

Finally, we sketch how one might introduce other constraints to study deviations from the classes of games above. We consider potential games, in which all differences in payoffs for unilaterally switching to another action can be expressed using a global potential function (Monderer and Shapley, 1996b). (See Section 1.2.3 for a definition.) The potential function

maps each action profile to a real number. We generate payoff matrices at random, and for each payoff matrix we also generate an associated potential function. We then modify the game so that the differences in payoffs conform to the potential function. This is tuned by a parameter $\xi \in [0, 1]$: when $\xi = 0$ there is no modification, and so the game is completely random, and when $\xi = 1$ the payoff matrix describes a perfect potential game.

More formally, denoting by $\Pi^{\star,R}$ and $\Pi^{\star,C}$ the modified payoff matrices, the new payoffs are

$$\begin{aligned} \Pi_{ij}^{\star,R} &= (1-\xi)\Pi_{ij}^{R} + \xi\left(\Pi_{1j}^{R} + V_{ij} - V_{1j}\right), \forall i \in 2,\ldots,N, \forall j \in 1,\ldots,N, \\ \Pi_{ji}^{\star,C} &= (1-\xi)\Pi_{ji}^{C} + \xi\left(\Pi_{1i}^{C} + V_{ij} - V_{i1}\right), \forall i \in 1,\ldots,N, \forall j \in 2,\ldots,N. \end{aligned} \tag{3.2}$$

For each pair (N, ξ), with $N \in [3, 10, 50]$ and $\xi \in [0.0, 0.1, \ldots, 0.5, \ldots, 0.9, 1.0]$, we generate 200 payoff matrices and associated potential matrices at random, and then we modify the payoffs using Eq. (3.2). For each payoff matrix we simulate each of the six learning algorithms starting from 10 different initial conditions. We measure the share of best reply cycles and the nonconvergence frequency of each algorithm. The results are reported in Figure 3.9.

As in Figure 3.5, the share of best reply cycles generally tracks the non-convergence frequency of the six algorithms. This is mostly true for $N = 50$.[33] The approximation is worse for $N = 3$, where it works well only for replicator dynamics. However, when $N = 3$ several algorithms have a nonconvergence frequency of about 10% for small values of ξ, and the nonconvergence frequency drops to 0% as $\xi = 1$ is approached. Differently from Figure 3.5, fictitious play behaves like the majority of the learning algorithms we consider. The striking exception here is EWA with noise. We conjecture that the structure of potential games and the effect of noise makes this learning algorithm "jump" between multiple pure strategy Nash equilibria.

For all values of N the share of best reply cycles drops to zero as ξ approaches 1.[34] However, it does so in a non-linear way that depends on N. For $N = 3$ small deviations from potential games imply a small share of best reply cycles, while for larger values of N the share of best reply cycles can be substantial. When $N = 3$ and $\xi = 0.8$, it is $\langle \mathcal{F}(\boldsymbol{v}) \rangle_{N,\xi} = 0.05$ only, but for $N = 50$ and $\xi = 0.8$ it is $\langle \mathcal{F}(\boldsymbol{v}) \rangle_{N,\xi} = 0.35$.

This suggests that in some situations small deviations from commonly studied classes of games may cause a significant increase in the share of best reply cycles, making them less stable under learning.

[33] Note that in this case the non-convergence frequency of fictitious play is slightly above zero, in contrast with the theorem by Monderer and Shapley (1996a). This minor discrepancy is simply due to a transient effect.

[34] This result is in line with Candogan et al. (2013), who showed that fictitious play and other related algorithms always converge as long as deviations from potential games are sufficiently small.

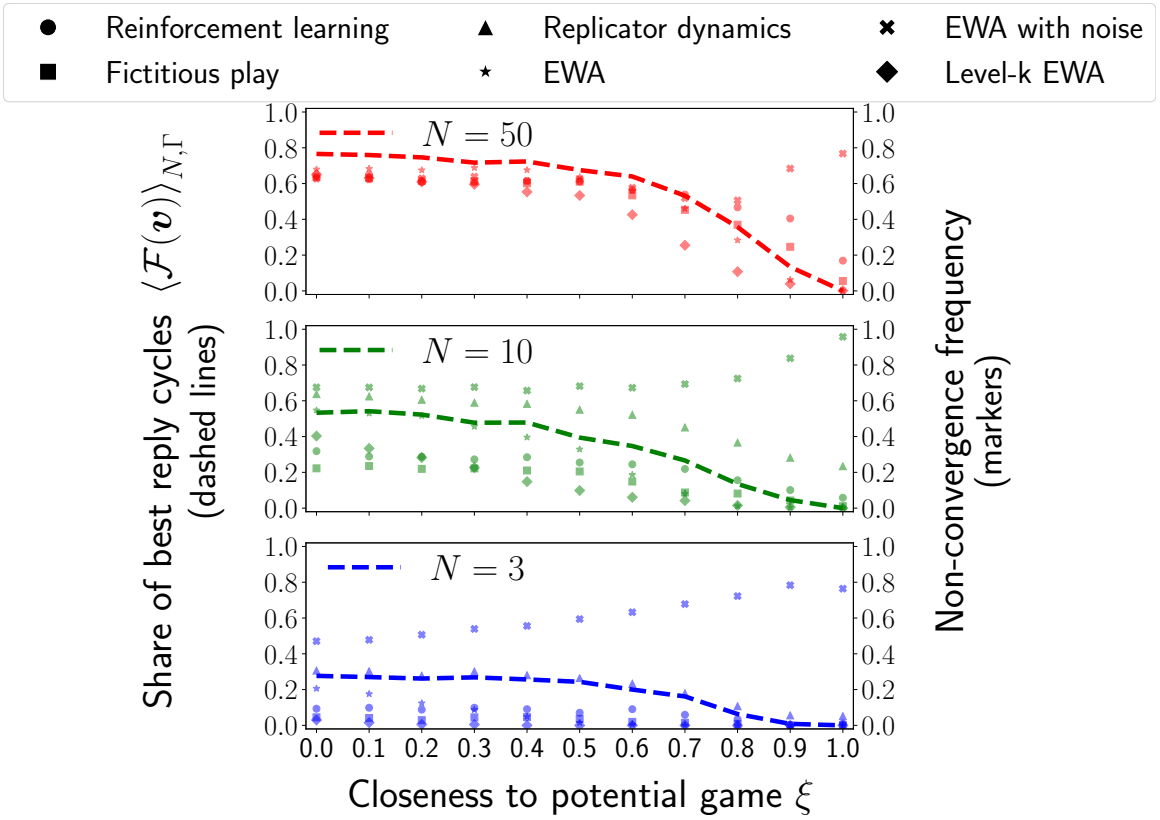


Figure 3.9: **How the share of best reply cycles predicts convergence as perfect potential games are approached.** Dashed lines are the average share of best reply cycles $\langle \mathcal{F}(\boldsymbol{v}) \rangle_{N,\xi}$ for those values of N and ξ. Here N is the number of actions and ξ is the closeness to potential games. Markers are the fraction of simulation runs in which the learning algorithms do not converge.

3.6 Analytical approach

For $\Gamma = 0$ (and $\xi = 0$) it is possible to derive analytically how the best reply structure varies with N. Our derivation follows the framework of statistical mechanics, suggesting that problems such as the existence of phase transitions in the best reply structure may be studied within our formalism. We define a *best reply configuration* as a unique set of best replies by both players to all possible actions of their opponent. Since the best reply matrix is a Boolean matrix, for any given N there is a countable number of possibilities. The total number of possible best reply configurations is N^{2N}. For uncorrelated payoffs, $\Gamma = 0$, all best reply configurations are equally likely. Therefore we can compute the frequency $\rho(\boldsymbol{v})$ for any set of attractors $\boldsymbol{v}$ by counting the number of best reply configurations leading to $\boldsymbol{v}$. In the jargon of statistical mechanics, this means that we are assuming a micro-canonical ensemble of games.

Once we have an expression for the frequency of each $\boldsymbol{v}$, this can be used to compute the ensemble average of the share of best reply cycles $\mathcal{F}$ for any given N (which we showed is

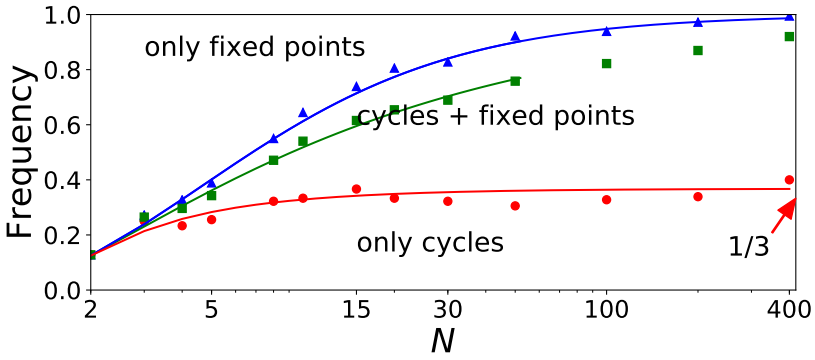


Figure 3.10: **Comparison of analytical predictions about best reply cycles to numerical simulations when $\Gamma = 0$.** Markers are numerical results and solid lines are analytical results. Red circles depict the frequency of randomly generated payoff matrices with no fixed points ($\mathcal{F}(\boldsymbol{v}) = 1$), blue triangles show the frequency with at least one cycle ($\mathcal{F}(\boldsymbol{v}) > 0$). The text in the figure refers to the area delimited by solid lines, e.g. "cycles + fixed points" means that the fraction of payoff matrices with both cycles and fixed points is the distance between the red and blue lines. Finally, green squares represent the average share of best reply cycles $\mathcal{F}_N$; this is discontinued at $N = 50$ due to excessive computational cost, see Section 3.6.2.

a good predictor for the non-convergence frequency of several learning algorithms in games with that value of N),

$$\mathcal{F}_N = \sum_{\boldsymbol{v}} \rho(\boldsymbol{v})\mathcal{F}(\boldsymbol{v}), \tag{3.3}$$

summing over all possible $\boldsymbol{v}$ s.t. $\sum_{k=1}^{N} n_k k \leq N$. It is also possible to calculate other quantities, including the fraction of payoff matrices without fixed points ($\mathcal{F}(\boldsymbol{v}) = 1$) and without cycles ($\mathcal{F}(\boldsymbol{v}) = 0$).

Before presenting the calculation of $\rho(\boldsymbol{v})$ in Section 3.6.1, and the other calculations in Section 3.6.2, we show the agreement between our analytical results and Monte Carlo simulations in Figure 3.10.

In this figure we analyze the best reply structure for increasing values of N. We report, from bottom to top, the fraction of payoff matrices with no fixed points, the average share of best reply cycles $\mathcal{F}_N$, and the fraction of games with at least one cycle. For instance, for $N = 30$, 36% of the payoff matrices have no fixed points, 84% have at least one cycle, (so 16% have no cycles, and 48% have a mixture of cycles and fixed points), with an average $\mathcal{F}_N = 0.70$. There is a very good agreement between analytical results (solid lines) and Monte Carlo sampling (markers). The fraction of games with cycles is an increasing function of N; it is computationally intractable to compute this for large N, but it seems to be tending to one. However, the fraction of games with no fixed points seems to reach a fixed value for $N \to \infty$. In Section 3.6.2 we show that this is approximated by $1/3$, in agreement with numerical simulations and close to the exact result $1/e$ (Goldberg et al., 1968).

3.6.1 Frequency of best reply vectors

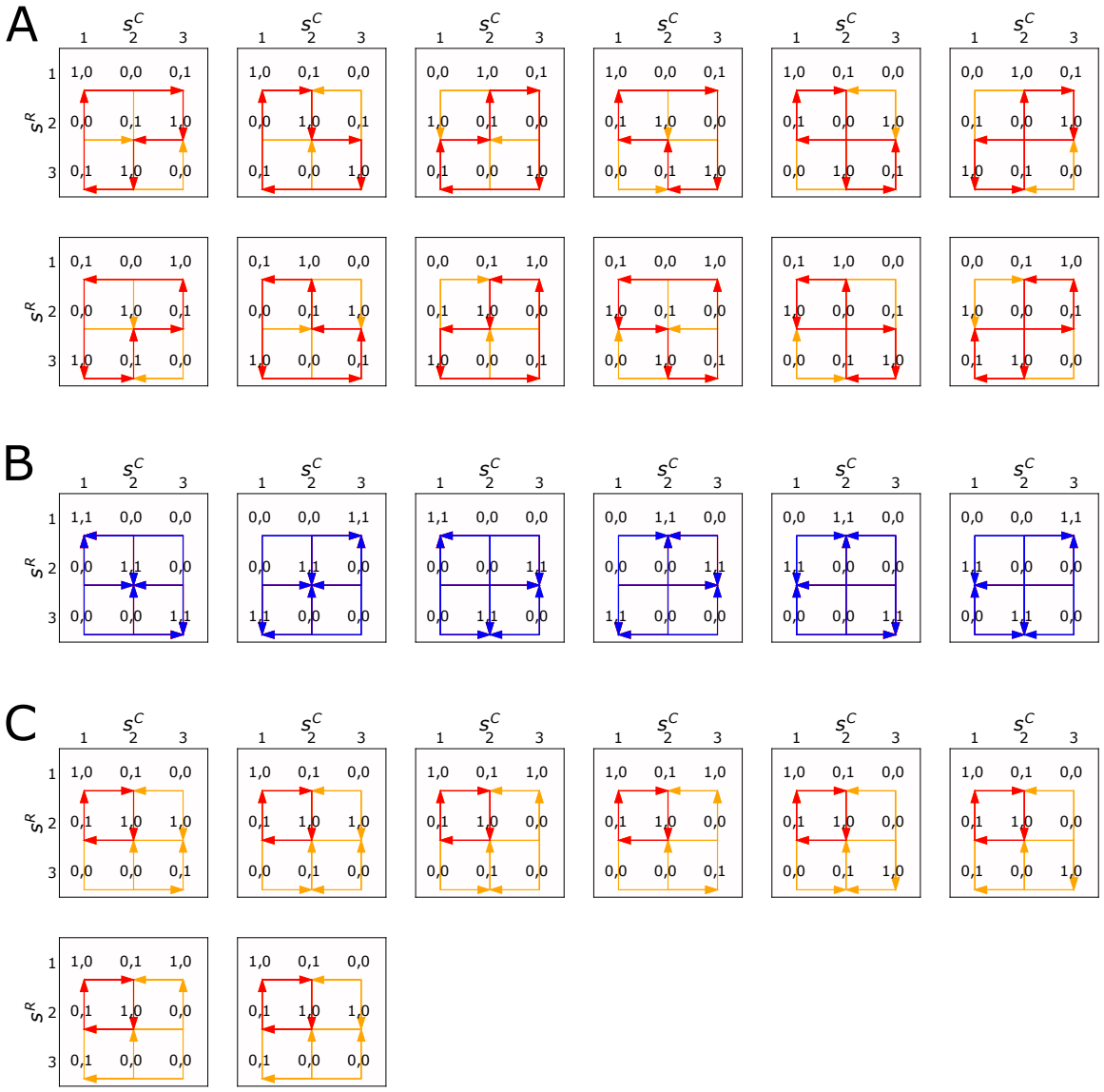


Figure 3.11: **Exhaustive account of best reply configurations with the same best reply vector.** (A) All possible $3! \ \ 2! \ = 12$ ways to combine 3 actions per player to form 3-cycles. Using Eq. (3.4), $f(3,3) = 12$. (B) All possible $3! = 6$ ways to combine 3 actions per player to form 3 fixed points. Using Eq. (3.4), $f(3,1)\frac{f(2,1)}{2}\frac{f(1,1)}{3} = 9\frac{4}{2}\frac{1}{3} = 6$. (C) All possible $3^2 - 1 = 8$ ways to choose the two free best replies when a 2-cycle is in place, so that the free best replies do not form a fixed point at $(3,3)$. Using Eq. (3.6), $g_3(1,0) = 8$.

We first discuss the count of the ways to form k-cycles and fixed points of best reply dynamics, and then we count the ways to place the *free best replies* (i.e. those that are not part of either cycles or fixed points). Finally we show how we combine these numbers together to obtain the count of best reply configurations that correspond to a specific set of attractors.

We start the count of k-cycles by example. In Figure 3.11A we exhaustively report all possible ways to form 3-cycles in a payoff matrix with $N = 3$, corresponding to all possible best reply sequences that form a closed loop. We arbitrarily start at $s^R = 1$ (because this is a cycle, the starting point does not matter), we look at the best reply by player Column, $s^C \in \{1, 2, 3\}$, and we connect $s^R = 1$ with s^C. In the top left panel, we connect $s^R = 1$ to $s^C = 3$. The first choice can be done in $k = 3$ ways. Once we have determined the first best reply by Column, we continue constructing the cycle by choosing a second best reply by Row. The second choice can only be done in $k - 1 = 2$ ways. In the top left panel, we connect $s^C = 3$ to $s^R = 2$. We then select a second best reply by Column. Again, we have $k - 1 = 2$ possibilities. In the top left panel, we connect $s^R = 2$ to $s^C = 2$. The third and last best replies for Row and Column are constrained, there is only one ($k - 2 = 1$) way to choose the remaining best reply. In the top left panel, we connect $s^C = 2$ to $s^R = 3$ and $s^R = 3$ to $s^C = 1$. We have $3 \cdot 2 \cdot 2 \cdot 1 \cdot 1 = 12$ ways to form 3-cycles with $n = 3$ available actions. Here n denotes the number of actions per player which are not already part of cycles or fixed points. In general n might be smaller than N, but in Figure 3.11A all actions are part of the 3-cycle, so $N = n = k = 3$.

It is possible to generalize this argument and to conclude that there are $k!(k-1)!$ ways to form k-cycles, once we determine which actions of players Row and Column are involved. Any k actions out of n can be chosen (by both players), so there are $\binom{n}{k}^2$ possibilities. We define

$$f(n,k) = \binom{n}{k}^2 k!(k-1)!, \tag{3.4}$$

with $2 \leq k \leq n$, as the count of the ways to have a k-cycle with n available actions per player. In the above example, $f(3,3) = 12$.

We now look at the ways to form fixed points, and we begin again by example. In Figure 3.11B we report all possible ways to form 3 fixed points in a payoff matrix with $N = 3$. Once we determine which actions are part of the fixed points (all, in this case), we form all possible combinations of fixed points by picking pairs of actions from the lists of available actions by both players. For convenience, we start again from $s^R = 1$. We form a fixed point by choosing any action $s^C \in \{1, 2, 3\}$, so that $\left(s^R, s^C\right)$ is a fixed point. In the left panel, we choose (1,1) as the first fixed point. We then consider $s^R = 2$. There are only two actions available from player Column to form a second fixed point. In the left panel, (2,2) is the second fixed point. Finally, for $s^R = 3$ only one action by Column is available. By process of elimination, in the left panel (3,3) is the third and last fixed point.

This example illustrates that the computation of the number of fixed points is very similar to the case of cycles, and indeed fixed points are just cycles of length one. In order

to get the number of ways to form fixed points, we can apply Eq. (3.4) iteratively, being careful to avoid the double, triple etc. counting of fixed points. We get

$$\prod_{j=1}^{n_1} \frac{f(n+1-j,1)}{j} \tag{3.5}$$

as the count of the ways to have n_1 fixed points with n available actions per player dividing by j to avoid multiple counting (see below). In the above example, $f(3,1)\frac{f(3,2)}{2}\frac{f(3,3)}{3} = 9\frac{4}{2}\frac{1}{3} = 6$.

We finally calculate the ways to place the free best replies, which are not part of either cycles or fixed points. We begin again by example. In Figure 3.11C we show payoff matrices with one free best reply per player. In the top left panel, the best reply of Row to Column playing $s^C = 3$ is $s^R = 2$; the best reply of Column to Row playing $s^R = 3$ is $s^C = 3$. The free best replies can be chosen freely, except for both of them to be action 3, in which case they would form another fixed point. In this example there are $3^2 - 1 = 8$ ways to choose free best replies so that they do not form other cycles or fixed points.

In general,

$$g_N(n,d) = N^{2n} - \sum_{k=1}^{n} f(n,k) g_N(n-k, d+1)/(d+1) \tag{3.6}$$

counts all possible ways to combine n free best replies in a $N \times N$ payoff matrix, so that they do not form other cycles or fixed points. We provide a more complete example for Eq. (3.6) below in this section. Note that N is a parameter and therefore is indicated as a subscript, while n is a recursion variable: even when the number of available actions n is smaller than N, the free best replies can be chosen out of all the N actions (see Figure 3.11C), in N^{2n} ways. The second term counts the "forbidden" combinations, i.e. the ones that form cycles or fixed points. This term has a recursive structure. It counts the number of ways to form each type of attractor, and then the number of ways not to have other attractors. d denotes the recursion depth. The division by $d+1$ is needed to prevent double, triple, etc. counting of attractors.

We now combine all the ways to have cycles, fixed points and free best replies to calculate the number of best reply configurations that correspond to a generic best reply vector $\boldsymbol{v} = (n_N, n_{N-1}, ...n_k, ...n_2, n_1)$. We denote by n_1 the number of fixed points and by n_k, with $2 \leq k \leq N$, the number of k-cycles. Of course $\boldsymbol{v}$ has to obey the obvious constraint that fixed points and k-cycles do not take up more than N actions: $\sum_{k=1}^{N} n_k k \leq N$. The frequency of the best reply vector $\boldsymbol{v}$ is

$$\rho(\boldsymbol{v}) = \left(\prod_{k=1}^{N} \prod_{j=1}^{n_k} \frac{f\left(N - \sum_{l=k+1}^{N} n_l l - (j-1)k, k\right)}{j} \right) g_N\left(N - \sum_{l=2}^{N} n_l l - n_1, 0\right) \Big/ \left(N^{2N}\right). \tag{3.7}$$

The first term with f counts all the ways to have k-cycles, by multiplying the counts for all values of k (first product) and for all k-cycles for a specific value of k (second product). Note that we progressively reduce the number of actions available to form k-cycles, as more and more actions become part of k-cycles (see below for an example that clarifies this point). If there are multiple k-cycles, $n_k > 1$, we divide the count by $j = 1, ...n_k$ so to avoid double, triple, etc. counting. The case $k = 1$ accounts for fixed points. The second term g_N counts all the ways to choose the remaining $N - \sum_{l=2}^{N} n_l l - n_1$ free best replies. The product of the three terms gives the number of best reply configurations that correspond to the best reply vector $\boldsymbol{v}$. We divide this number by the possible configurations N^{2N} and we obtain the frequency $\rho(\boldsymbol{v})$.

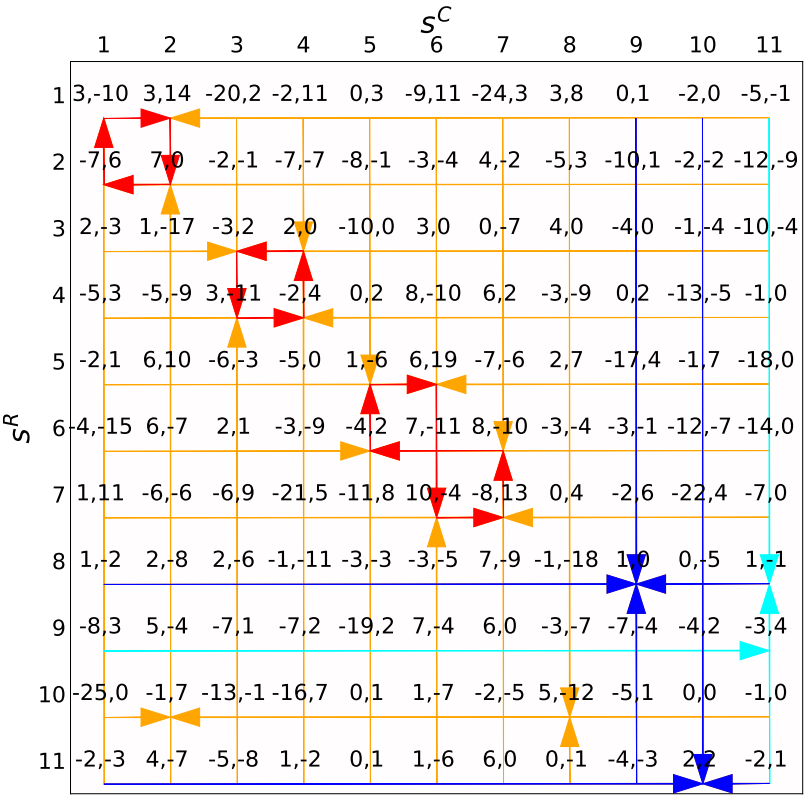


Figure 3.12: **Payoff matrix with $N = 11$ that is used to illustrate the calculation of the frequency of the best reply vectors.** The best reply vector is $\boldsymbol{v} = (0, 0, 0, 0, 0, 0, 0, 0, 1, 2, 2)$, with $n_3 = 1$, $n_2 = 2$, $n_1 = 2$ and $n_k = 0$, if $k > 3$. It is $\rho(0, 0, 0, 0, 0, 0, 0, 0, 1, 2, 2) = 1.44 \cdot 10^{-6}$, so this best reply vector is very infrequent.

As an example, we calculate the number of best reply configurations with the same best reply vector as in Figure 3.12. We start counting the ways to form 3-cycles. We can choose any 3 actions out of 11 for both players to be part of a 3-cycle, meaning that there are $\binom{11}{3}^2$ possibilities. Once we have selected 3 actions per player, we can obtain 12 cycles for each choice by choosing $3!2! = 12$ sequences of actions. So the number of ways to form 3-cycles is

$f(11,3)$. The same reasoning applies to the two 2-cycles, except that there are only 8 and 6 actions per player still available and that the count of the ways to have 2-cycles needs to be divided by 2 in order to avoid double counting. So we multiply $f(11,3)$ by $f(8,2)f(6,2)/2$. The number of best reply configurations with 2 fixed points in the remaining 4 actions can be calculated similarly: each player can choose the first fixed point out of 4 actions, and the second out of 3, but we have to consider double counting. So $f(4,1)f(3,1)/2$ gives the ways to form the two fixed points out of the 4 remaining actions. We are left with 2 actions per player that are not part of cycles or fixed points. There are 11^4 ways to choose the free best replies, but we have to exclude the cases in which they would form another 2-cycle or one or more fixed points. There are 2 ways they could form a 2-cycle ($f(2,2)$), and 4 ways they could form 1 fixed point ($f(2,1)$). But for each of the latter we have to consider all compatible configurations, i.e. calculate $g_{11}(1,1)$: there are 11^2 ways to choose the free best replies, minus the way in which this choice would form another fixed point (divided by 2, to account for the situation with two fixed points). In summary, the number of best reply configurations is given by

$$\rho(0,0,0,0,0,0,0,0,0,1,2,2) = f(11,3)f(8,2)\frac{f(6,2)}{2}f(4,1)\frac{f(3,1)}{2}g_{11}(2,0)/(11^{22}), \tag{3.8}$$

with $f(11,3) = \binom{11}{3}^2 3\cdot 2\cdot 2$, $f(8,2) = \binom{8}{2}^2 2$, $f(6,2) = \binom{6}{2}^2 2$, $f(4,1) = \binom{4}{1}^2 1$, $f(3,1) = \binom{3}{1}^2 1$ and $g_{11}(2,0) = 11^4 - 2 - 4\cdot g_{11}(1,1)$, with $g_{11}(1,1) = 11^2 - 1/2$.

The explicit computation of the frequency gives $\rho(0,0,0,0,0,0,0,0,1,2,2) = 1.44\cdot 10^{-6}$, so the best reply vector in Figure 3.12 is very infrequent. For $N = 11$, the most common best reply vectors are instead:

$$\begin{aligned}
&\rho(0,0,0,0,0,0,0,0,0,0,1) = 0.17,\\
&\rho(0,0,0,0,0,0,0,0,0,0,2) = 0.14,\\
&\rho(0,0,0,0,0,0,0,0,0,1,0) = 0.14,\\
&\rho(0,0,0,0,0,0,0,0,0,1,1) = 0.13,\\
&\rho(0,0,0,0,0,0,0,0,1,0,0) = 0.09.
\end{aligned} \tag{3.9}$$

For $N = 20$, the most common best reply vectors are:

$$\begin{aligned}
&\rho(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1) = 0.10,\\
&\rho(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1) = 0.10,\\
&\rho(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,2) = 0.09,\\
&\rho(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0) = 0.09,\\
&\rho(0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0) = 0.07.
\end{aligned} \tag{3.10}$$

We observe that k-cycles with high values of k are never really frequent; the frequency of any specific best reply vector decreases with N (because there are many more best reply vectors with positive frequency); the best reply vectors with cycles become more frequent as N increases, consistently with Figure 3.10. Note that an accurate numerical estimate of the most common best reply vectors might be challenging due to the extremely high number of best reply configurations: the analytical results make it possible to obtain exact estimates.

3.6.2 Other combinatorial calculations

So far we have provided an analytical expression to calculate the frequency of a specific best reply vector. In this section we obtain equations for the frequency of payoff matrices with at least one fixed point or one cycle of any specific length, and then for the frequency of payoff matrices with at least one cycle of any length. We finally derive an expression for the frequency of k-cycles in the $N \to \infty$ limit.

These expressions are useful because it is computationally very expensive to calculate the frequency of all best reply vectors and then consider the ensemble average. Indeed, in Figure 3.10 the analytical line with the share of best reply cycles averaged over the ensemble (middle green line, $\mathcal{F}_N$) stops at $N = 50$. On the contrary, the analytical lines for the fraction of payoff matrices with at least one cycle (top blue line, $\mathcal{F}(\boldsymbol{v}) > 0$) and with no fixed points (bottom red line, $\mathcal{F}(\boldsymbol{v}) = 1$) continue up to $N = 400$. This is due to the fact that to compute the middle line we need to explicitly calculate the frequency of all best reply vectors, whereas to compute the top and bottom lines we use the expressions derived in this section.

Define

$$h_N(n,k,d) = f(n,k)\left[N^{2(n-k)} - \frac{h_N(n-k,k,d+1)}{d+2}\right]. \tag{3.11}$$

h_N counts the number of configurations with at least one k-cycle in a $N \times N$ payoff matrix, with n actions that are not already part of other k-cycles, at recursion depth d. The reasoning is similar to that in the previous section. Consider for instance the calculation of the number of 2-cycles in a 4×4 payoff matrix: $N = n = 4, k = 2, d = 0$. By using Eq. (3.11), $h_4(4,2,0) = f(4,2)\left[4^{2\cdot 2} - h_4(2,2,1)/2\right]$, where $h_4(2,2,1) = f(2,2)\left[4^0\right] = 2$. There is a number $f(4,2)$ of 2-cycles, and for each of these there are 4^4 ways to place the two remaining best replies of the players. But if those are combined so that they form another 2-cycle, we would count 2-cycles twice, so we need to remove one best reply configuration from the count.

We use the shorthand

$$\rho(N,k) = \frac{h_N(N,k,0)}{N^{2N}} \tag{3.12}$$

for the fraction of $N \times N$ payoff matrices with at least one k-cycle. Because a fixed point is a cycle of length one, Eq. (3.12) can be used to calculate the number of payoff matrices with at least one fixed point, and

$$\rho_N(n_1 = 0) = 1 - \frac{h(N,1)}{N^{2N}} \tag{3.13}$$

is the fraction of payoff matrices with no fixed points. Eq. (3.13) has been used for the bottom red analytical line in Figure 3.10. Best reply dynamics never converges to a fixed point in these games, and other learning algorithms are very unlikely to converge as well (see Figure 3.2). Therefore, $\rho_N(n_1 = 0)$ is a lower bound for the frequency of non-convergence in generic games with N actions.

Now define

$$h'_N(n,d) = \sum_{k=2}^{n} f(n,k) \left[N^{2(n-k)} - \frac{h_N(n-k,k,d+1)}{d+2} \right]. \tag{3.14}$$

This expression is analogous to Eq. (3.11), but it considers k-cycles of any length (except $k = 1$), as opposed to k-cycles of a specific length. Indeed, we sum over all possible values of k, and the term with the double counting also considers cycles of any length. The fraction of configurations with at least one cycle is

$$\rho_N \left(\sum_{k=2}^{N} n_k > 0 \right) = \frac{h'_N(N,0)}{N^{2N}}, \tag{3.15}$$

and this expression has been used for the top blue analytical line in Figure 3.10. It represents an upper bound for the frequency of non-convergence in generic games with N actions, because the lack of best reply cycles implies convergence in most cases.

Note that $\sum_{k=2}^{N} \rho(N,k)$ sums to more than N^{2N}, because several best reply configurations have multiple cycles of different length. On the contrary, $h'_N(N,0)$ is always less than N^{2N}, because some configurations have cycles but no fixed points. Had we started the summation in Eq. (3.14) from $k = 1$, the count would sum exactly to N^{2N}, because all configurations have at least one cycle or one fixed point.

We finally calculate the asymptotic frequency of attractors in the limit $N \to \infty$. Eq. (3.11) can be used in this limit to calculate analytically the absolute and relative frequencies of payoff matrices with at least one k-cycle or fixed point. We make the following ansatz:

$$\lim_{N\to\infty} \frac{h_N(N,k,0)}{N^{2N}} \approx \frac{h_N(N-k,k,1)}{(N-k)^{2(N-k)}}. \tag{3.16}$$

We are making two approximations whose validity will be verified ex-post. First, the frequency of k-cycles reaches a fixed point as $N \to \infty$. Second, the functional form of $h_N(n,k,0)$ is very similar to that of $h_N(n,k,1)$. We know that this is not the case, as the

term used to avoid multiple counting – namely $h_N(n-k,k,d+1)$ – is divided by 2 for $d=0$ and by 3 for $d=1$. The approximation becomes exact only for $d \to \infty$ (because $1/d$ and $1/(d+1)$ are very similar), but the quantity we are interested into has $d=0$.

We can write

$$\frac{h_N(N,k,0)}{N^{2N}} = \frac{N^2(N-1)^2...(N-k+1)^2}{(k!)^2} k!(k-1)! \frac{\left[N^{2(N-k)} - \frac{h_N(N-k,k,1)}{2}\right]}{N^{2N}}. \tag{3.17}$$

By applying the ansatz in Eq. (3.16) and after some algebra we obtain

$$\lim_{N\to\infty} \frac{h_N(N,k,0)}{N^{2N}} := \rho(k) = \frac{1}{(k!)^2} k!(k-1)! \left(1-\rho(k)/2\right), \tag{3.18}$$

which can be solved self-consistently to yield

$$\rho(k) = \frac{2}{2k+1}. \tag{3.19}$$

So for $N \to \infty$, fixed points appear in 2/3 of the payoff matrices, 2-cycles appear in 2/5 of the payoff matrices, 3-cycles in 2/7, 4-cycles in 2/9, etc. Eq. (3.19) has been used to calculate the asymptotic frequency of configurations with no fixed points (1/3) in Figure 3.10. We can easily obtain the relative frequencies (with respect to fixed points):

$$\frac{\rho(k)}{\rho(1)} = \frac{3}{2k+1}, \tag{3.20}$$

so 2-cycles appear 3/5 as often as fixed points, 3-cycles appear 3/7 as often, 4-cycles 3/9 as often, 5-cycles 3/11 as often, etc.

In Figure 3.13 we report the frequency of k-cycles, as calculated using Eq. (3.17), as a function of the number of available actions N. There is a good correspondence between the asymptotic behavior in Eq. (3.19) and the explicit computation up to $N = 400$, at least for the smallest values of k (excluding the fixed points).

3.7 Perspectives

In collaboration with Torsten Heinrich, Luca Mungo, Bassel Tarbush, Samuel Wiese, Yoojin Yang and Doyne Farmer, we have been extending the best reply framework to multiplayer network games. In that setting, the best reply structure still predicts the convergence frequency of the learning algorithms considered in this thesis. In multiplayer games in which each player is connected to all other players, the share of best reply cycles increases with the number of players. The share of best reply cycles also increases with the average degree of the interaction network, for a given number of players. We conjecture that an approximate scaling equivalence between the number of actions, the number of players and

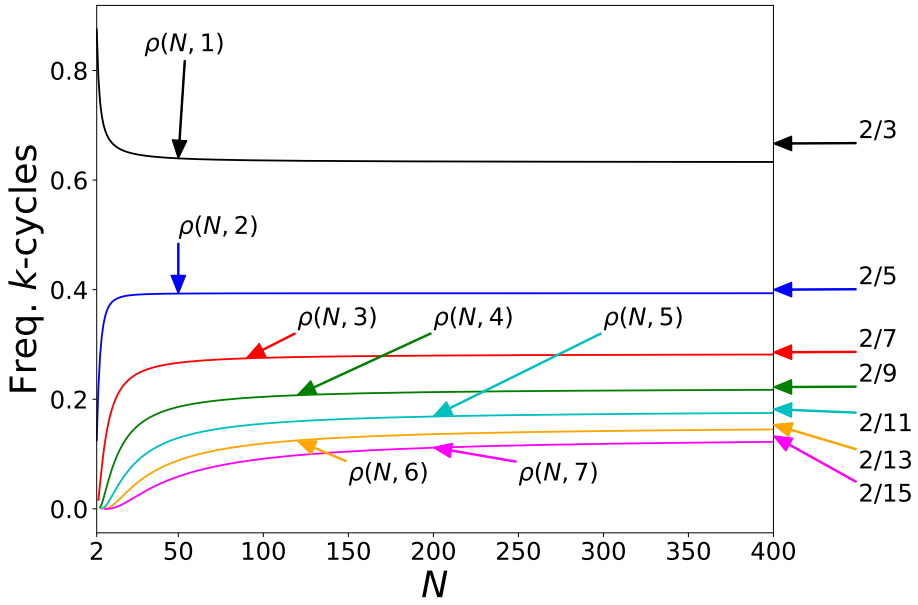


Figure 3.13: **Frequency of k-cycles, $\rho(N,k)$, as a function of the number of actions N.** The numbers annotated on the right are the asymptotic frequencies of k-cycles, as calculated using Eq. (3.19). The approximations tend to slightly overestimate the frequency, at least up to $N = 400$, even more so for larger values of k. The exception is the fixed points, in which the approximation tends to be worse.

the connectivity pattern determines the average share of best reply cycles. More precisely, our conjecture is that a game with M players and N actions approximately has the same share of best reply cycles as a game with 2 players and N^{M-1} actions. Moreover, a game with M players, N actions and d-regular connectivity pattern has approximately the same share of best reply cycles as a $(M + d + 1)/2$ player, N action game with fully connected interaction network. These preliminary result support the generality of the theory presented in this thesis.

In collaboration with Luca Mungo, Torsten Heinrich and Doyne Farmer, we have studied how well an alternative to best reply dynamics predicts the convergence frequency of the learning algorithms studied in this thesis. This alternative is a version of "better-reply dynamics" modeled as a Markov Chain. The state space of the Markov Chain corresponds to the possible actions of the players and the transition probabilities depend on the payoff differences for switching from one action to another. Best reply dynamics is obtained as a special case of this algorithm in which the transition probability to the best performing action is one and all other transition probabilities are zero. What this generalized algorithm captures is that in some cases there may not be a clear best reply, in the sense that two actions could yield a very similar payoff. Although this modification predicts the convergence frequency of learning algorithms better in these specific cases, it does not yield substantial improvements over the ensemble.

A first, technical, direction for future work is understanding more in detail why the predictions of the best reply formalism are so good. To understand this, a starting point would be to explicitly calculate the basin of attraction of some learning dynamics in the simplest Boolean best reply payoff matrices, such as the one given in Figure 3.1A, and formalize the intuition presented in Section 3.3. The next step could be to explicitly treat the best reply payoff matrix as a first order approximation, then study the behavior as one moves along a continuous path through the family of payoff matrices it is associated with. One could use this method to study the factors that cause deviations, such as the appearance and disappearance of mixed equilibria or better replies. This could potentially lead to a kind of perturbation expansion for understanding the convergence of learning algorithms in normal form games.

Our work can also be extended in more conceptual directions. One idea could be to develop on the constraints that define "real-world games". In Section 3.5 we have considered a constraint that parameterizes the extent to which differences in payoff conform to a potential function. Previous work has identified this constraint as relevant for many real-world situations, such as those involving congestion. We could exhaustively characterize properties of the best reply structure for various constraints. This naturally brings up extending our formalism to extensive-form games, given that strategic sequencing of actions is pervasive in the real world.

Moving away from pure game theory, there are several problems in economics that could benefit from our approach. For example, there is a large literature about convergence of dynamical processes to Walrasian equilibria. While it is possible to prove convergence in some cases, it is also possible to devise markets in which convergence does not occur (Saari, 1985). A priori, it is not clear whether real markets belong to one class or another. It would be interesting to assess whether convergence is "typical", in the sense used in this thesis, over market structures. Another problem for which our approach could be useful is convergence of learning to rational expectations equilibria (Evans and Honkapohja, 2012).

A more interdisciplinary direction would be to more closely relate our approach to statistical mechanics. In preliminary work, we have shown that the best reply structure can be expressed as a special type of Boolean network. Using results from random matrix theory for the adjacency matrix of the Boolean network we may prove additional properties on the prevalence of best reply cycles or pure equilibria.

Finally, it would be interesting to relate our results to the stability of ecosystems. An exact mapping exists between replicator dynamics and Lotka-Volterra equations (Hofbauer and Sigmund, 1998). It would be interesting to explore the equivalent mapping between the topological properties of the game—the best reply structure—and the topological properties

of the ecosystem, given by the structure of the corresponding food web. Our preliminary investigations show that loops in food webs are a sufficient but not necessary condition for best reply cycles in the corresponding payoff matrix. As mentioned in Section 3.2, the theory of qualitative stability in ecology has attempted to characterize stability of ecosystems based on the topological properties of the food web only. This theory did not develop much because some counterexamples showed that in some cases it failed predicting stability. Some counterexamples also show that the best reply structure can fail predicting convergence, yet over the ensemble it is a very accurate predictor given its simplicity. Similarly, it may be possible to quantify how often focusing on the topological structure of the food web is enough to assess stability of the ecosystem.

3.A Details of the simulation protocol

We describe here in detail the convergence criteria and the random generation of the games.[35] We need to simulate very different learning algorithms over potentially high-dimensional random games and to identify the simulation runs that converge to equilibrium. This leads to unavoidable arbitrary choices in the specification of the learning algorithms, the value of the parameters and the criteria that determine convergence. We have experimented a lot of combinations of design choices, and the overall picture is robust to the specific implementation for all sensible parameterizations.

3.A.1 Reinforcement learning

This algorithm is the deterministic Bush-Mosteller reinforcement learning specification of Macy and Flache (2002) and Galla and Farmer (2013), as explained in Section 1.3.2.

Convergence criteria. In Figure 3.14 we show instances of both converging and non-converging simulation runs. As is clear from the logarithmic plots in the bottom panels, no components of the mixed strategy vector ever reach a fixed point within simulation time. The reason is simple: Eqs. (1.17) do not have a memory loss term, so the probability for unsuccessful strategies keeps decreasing over time. Only numeric approximations would yield a fixed point, but under most parameter settings the Bush-Mosteller dynamics takes very long to reach the machine precision boundary.

Therefore, we choose a simple heuristic to determine if the learning dynamics has reached a fixed point:

1. Only consider the last 20% time steps.

[35] Our implementation in `Python` is available at `https://dx.doi.org/10.5281/zenodo.1419542`.

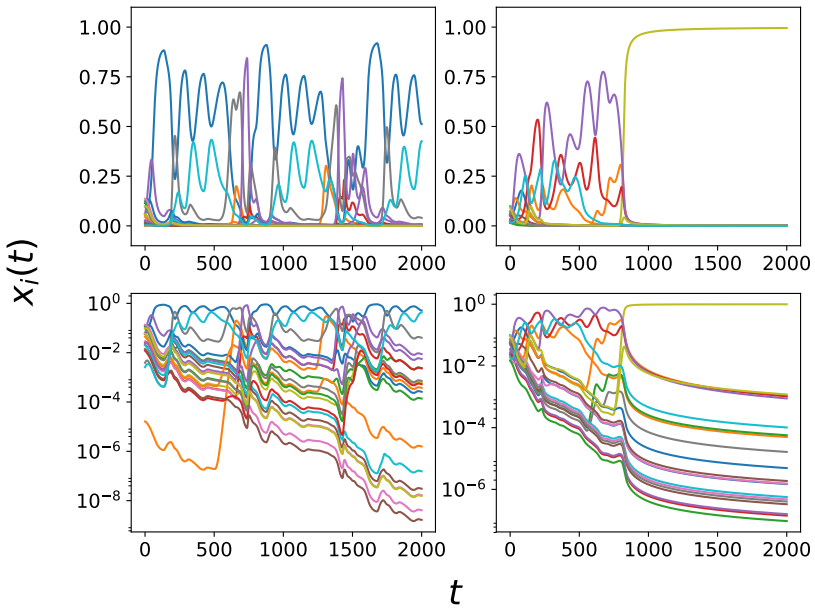


Figure 3.14: **Instances of simulation runs of the Bush-Mosteller reinforcement learning algorithm** with $N = 20$. Each line is a component of the mixed strategy vector of Row (not all components are visible, as they overlap). Left panels: non-converging simulation run. Right panels: converging simulation run. Top panels: linear scale. Bottom panels: logarithmic scale on the vertical axis. Only the first 2000 out of 5000 time steps are shown.

2. Only keep the actions that have been played with a probability larger than 0.05 (averaged over the time interval above).

3. If the average standard deviation (i.e. calculated over the time interval for each selected action and averaged over the selected actions) is larger than 0.01, identify the simulation run as non-convergent. Otherwise, if a component of the mixed strategy vector is larger than 0.98 identify the simulation run as convergent to a pure Nash equilibrium; if no component is larger than 0.98, consider the simulation run has having reached convergence to a mixed Nash equilibrium.

We experimented with slightly different specifications, with no significant effects on the results.

Parameter values. We do not observe much sensitivity to the parameter values. We perform the simulations with $\alpha = 0.2$ and $\beta = 0.5$, which avoid numerical problems. (For example, when β is too large the learning dynamics overshoots the strategy space boundaries.) We simulate the learning dynamics by iterating Eqs. (1.17) for 5000 time steps.[36]

[36]For $N = 5$, numeric approximations make the dynamics lose normalization after 2000 time steps. In this case we simulate for 2000 time steps only.

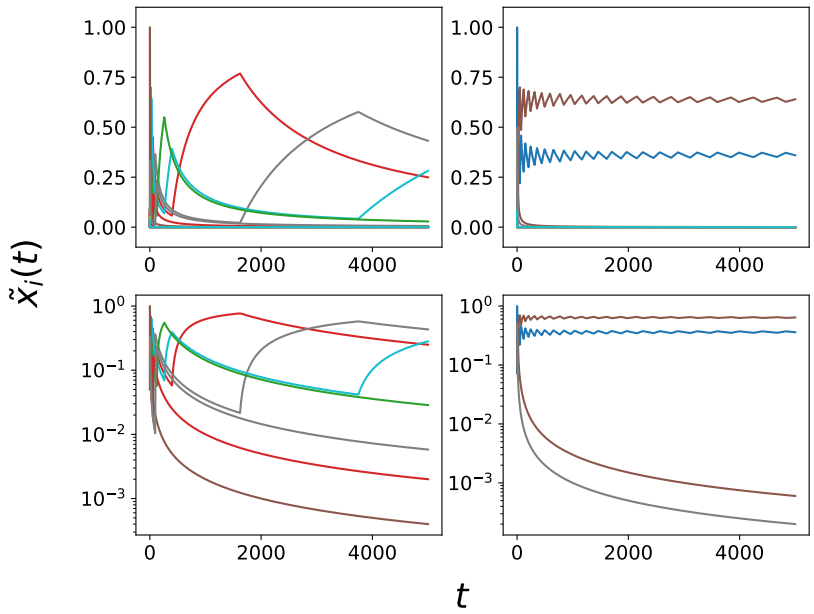


Figure 3.15: **Instances of simulation runs of fictitious play** with $N = 20$. Each line is a component of the mixed strategy vector of Row (not all components are visible, as they overlap). Left panels: non-converging simulation run. Right panels: converging simulation run. Top panels: linear scale. Bottom panels: logarithmic scale on the vertical axis.

3.A.2 Fictitious play

We study the standard fictitious play algorithm, as described in Section 1.3.3.

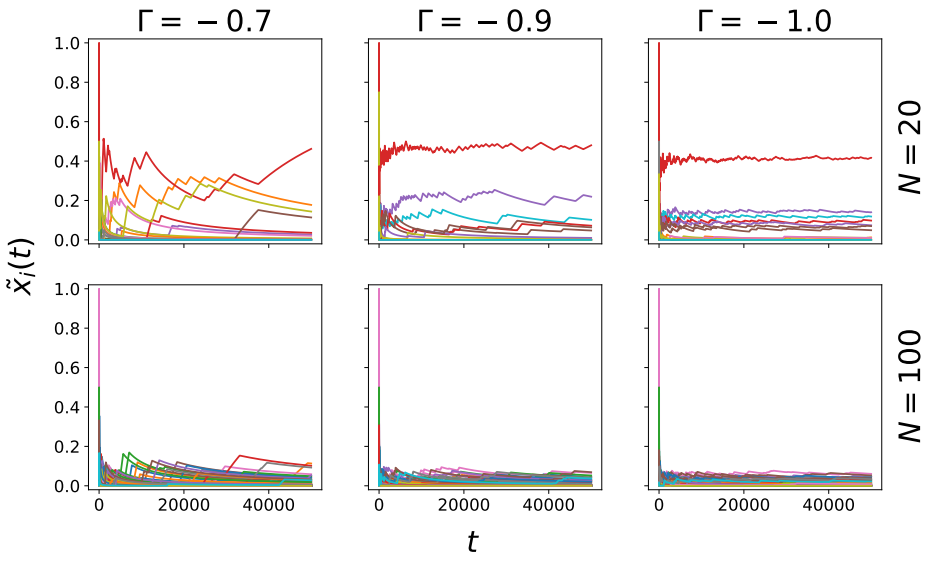


Figure 3.16: **Effect of negative payoff correlation on fictitious play** for some values of N and Γ. Only the simulation run with $\Gamma = -0.7$ and $N = 20$ is identified as non-convergent, all other simulation runs are identified as convergent. We show here 50000 time steps, although we only use 5000 iterations in the numerical analysis unless $N \geq 50$ (this is enough to measure the convergence rate almost as accurately).

Convergence criteria. We look at the convergence of the estimated mixed strategy vectors at time t, $\tilde{x}_i(t)$ and $\tilde{y}_j(t)$. This is commonly used to evaluate the convergence of fictitious play in games with no pure equilibria, as in this situation the actions actually taken cannot converge to a fixed point (Fudenberg and Levine, 1998). As it is clear from Figure 3.15, the behavior of fictitious play is very similar to that of Bush-Mosteller dynamics. Therefore, we use the same convergence criteria. A minor difference is that we identify a simulation run as having reached a mixed Nash equilibrium if no component of the mixed strategy vector is larger than 0.99 (instead of 0.98). The reason for this is that fictitious play reaches a pure Nash equilibrium in much shorter time than reinforcement learning, and so we can exclude transient behavior. Note that changing the expected strategies takes more and more time as t increases. In a certain sense, the behavior of the players becomes more set, as they need more sampling evidence to change their expectations.

Parameter values. Fictitious play has no parameters. We only need to choose the maximum number of iterations, which we take as 5000. For payoff matrices with $N \geq 50$ we consider longer time series (50000 time steps).

Note on negative correlation. It is well-known that fictitious play always converges to mixed strategy Nash equilibria in two-player zero-sum games. However, Robinson (1951) proved an *asymptotic* result that is valid only for $t \to \infty$ and $\Gamma = -1$. It is not clear: (i) how long it takes for the dynamics to come arbitrarily close to an equilibrium; (ii) what happens when $\Gamma > -1$ (still with anti-correlated payoffs, but not with opposite payoffs); (iii) how to measure convergence numerically. In Figure 3.16 we show the fictitious play dynamics for some values of Γ and N. The general message is that for more negative Γ and for larger N convergence becomes more likely, as only the $\Gamma = -0.7$, $N = 20$ simulation run is identified as non-convergent by the criteria above. However, convergence is never perfect in finite simulation time, and therefore the quantitative results depend on the convergence criteria.

For example, the fact that larger N makes convergence more likely is due to the fact that the dynamics moves closer to the center of the strategy space than with smaller N. So the trajectories are closer to the mixed Nash equilibrium in *absolute* terms, and the simulation run is identified as convergent even if some components of the mixed strategy vector are still changing by a small amount (in absolute terms). It is not clear if and how one should correct for this, for example making the threshold for the standard deviation in the convergence criteria depend on N, e.g. $0.01/N$. This would measure convergence in *relative* terms. In addition, the concentration of the trajectories in the center of the strategy space

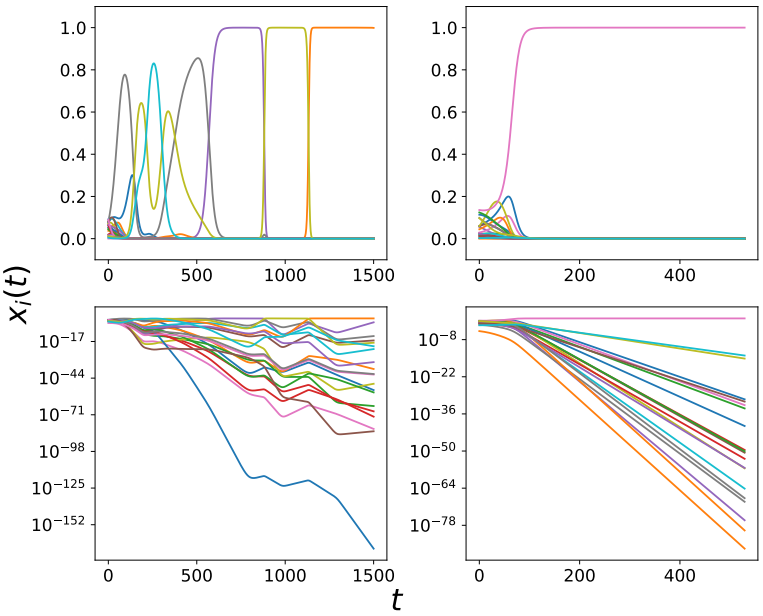


Figure 3.17: **Instances of simulation runs of replicator dynamics** with $N = 20$. Each line is a component of the mixed strategy vector of Row (not all components are visible, as they overlap). Left panels: non-converging simulation run. Right panels: converging simulation run. Top panels: linear scale. Bottom panels: logarithmic scale on the vertical axis. The maximum simulation time (1500 or 500 time steps) is endogenously determined by the first component of the mixed strategy vector hitting the machine precision boundary.

only occurs for substantially negative Γ, making it necessary to impose the dependence on N in an ad-hoc way. Given these problems, we think that the current convergence criteria are the most consistent choice.

3.A.3 Replicator dynamics

We use the discrete version of replicator dynamics defined in Section 1.3.4.

Convergence criteria. In Figure 3.17 we can see the technical problems associated with simulating the replicator dynamics. First, because only strict Nash equilibria are stable, all stable fixed points sit at the boundaries of the strategy space and cannot be reached in finite simulation time. Second, the period of cycles increases over time (due to the infinite memory of the replicator equations), and even unstable dynamics drifts towards the edges of the strategy space.

Third, while in the cases of Bush-Mosteller reinforcement learning and fictitious play the components of the mixed strategy vector were changing by relatively few orders of magnitude, the functional form of the replicator dynamics (1.28) implies an exponential change, as can be seen from the straight lines in the panels with logarithmic scale in Figure 3.17. Therefore, the map (1.28) can be reliably simulated only for a limited *confidence time*

interval: we stop the simulation run as soon as one component x_i or y_j reaches the machine precision limits.

This precaution is necessary because, if the dynamics is following a cycle, a certain action may not be played for a long time interval, with its probability decreasing over time. At some point, it may become convenient for the player to choose that action again, so the probability would start increasing again. But if the probability had hit the precision limits of the computer beforehand, it would be stuck at zero, falsely identifying the simulation run as having reached a fixed point.

Another problem concerns rounding approximations, which imply that normalization may be lost. To avoid this, we renormalize the mixed strategy vectors every time step.

With the parameter δ we choose, the confidence time interval is on average of the order of 1000 time steps (but can vary considerably, as can be seen in Figure 3.17). We could use the same convergence criteria as for Bush-Mosteller dynamics and fictitious play, but the short simulation time and the shape of the cycles – in linear scale, the dynamics is constant for a long time, and then suddenly changes – suggest to use a different heuristic. We check whether in the last 20% of the time steps the probabilities of the most used action for both players are monotonically increasing, while all other probabilities are monotonically decreasing. In other words

1. Only consider the last 20% time steps.
2. For each player, find the action with the highest probability, and verify whether this probability has been increasing for the full time interval.
3. Check that the probabilities for all other actions have been decreasing.
4. If conditions 2-3 are satisfied for both players, identify the simulation run as convergent.

These criteria simply reflect what we observe in Figure 3.17. While we cannot conclude that this heuristic works in general, a direct inspection of over 100 simulation runs for several values of N confirms that convergence to pure strategy Nash equilibria or failure to converge has been correctly identified in the vast majority of cases.

Finally, we would like to add a word of caution on the seemingly stronger instability of replicator dynamics as compared to other learning algorithms. Because of infinite memory and depending on the initial condition, it might take long to "find" a pure strategy Nash equilibrium, meaning that the replicator dynamics can hit the machine precision limits first, when it is still in a "transient". In other words, it may not be in the basin of attraction determined by a best reply cycle, but it may also have not reached a pure strategy Nash

equilibrium within the confidence time interval. This is especially the case for large payoff matrices, $N \geq 50$.

Parameter values. We simulate the replicator dynamics by setting the parameter $\delta = 0.1$, and a simulation time of 3000 time steps maximum. However, as discussed before, the simulation time is typically shorter and determined by the first strategy hitting the machine precision boundary.

3.A.4 Experience-Weighted Attraction

Experience-Weighted Attraction (EWA) is defined in Section 1.3.5.

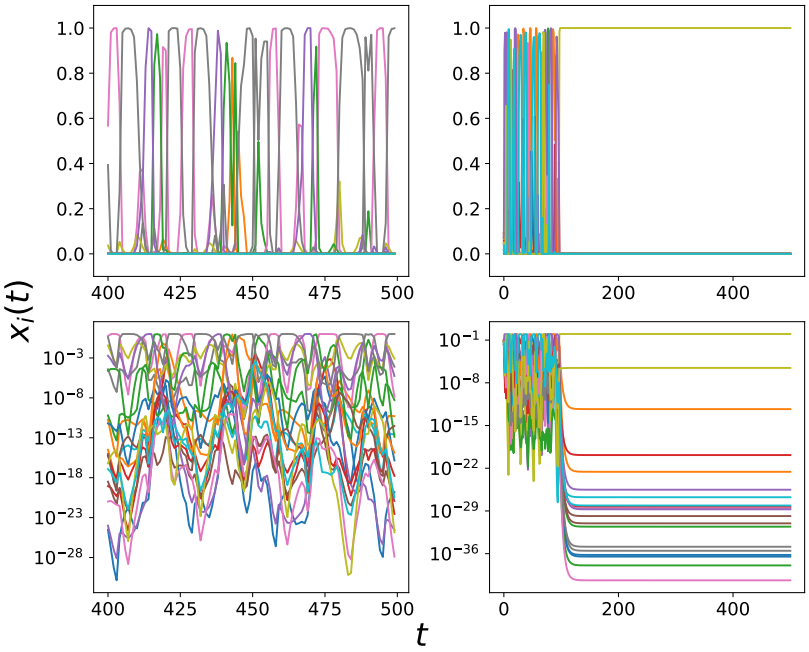


Figure 3.18: **Instances of simulation runs of Experience-Weighted Attraction** with $N = 20$. Each line is a component of the mixed strategy vector of Row (not all components are visible, as they overlap). Left panels: non-converging simulation run. The range of the horizontal axes has been chosen to ease visualization of the dynamics. Right panels: converging simulation run. Top panels: linear scale. Bottom panels: logarithmic scale on the vertical axis.

Convergence criteria. Consider Figure 3.18, right panels. All components of the EWA dynamical system reach a fixed point, differently from the other learning algorithms, so it is easier to identify convergence. We run the EWA dynamics for 500 time steps and we consider the last 20% time steps to determine convergence. With the parameter values we choose for α, β, κ and δ, the transient is usually of the order of 100 time steps, so 500 steps is enough to identify convergence. We then check that the average variance of the logarithms of the components of the mixed strategy vectors does not exceed a certain (very

small) threshold. We look at the logarithms because the probabilities following the EWA dynamics vary on an exponential scale and can be of the order of, e.g., 10^{-100}. In formula, if $1/N \sum_{i=1}^{N} 5/T \sum_{t=4/5T}^{T} (\log x_i(t))^2 > 10^{-2}$ or $1/N \sum_{j=1}^{N} 5/T \sum_{t=4/5T}^{T} (\log y_j(t))^2 > 10^{-2}$, with $T = 500$, we identify the simulation run as non-convergent. If convergence occurs to a mixed strategy vector in which a component exceeds 0.95, the simulation run is identified as convergent to a pure strategy equilibrium; otherwise, to a mixed strategy equilibrium.[37]

Parameter values. EWA has two main advantages from a computational point of view. First, if the memory loss parameter is positive ($\alpha > 0$), all stable attractors of the EWA system lie *within* the strategy space and not on the boundaries. This means that no actions are ever given null or unit probability and makes it possible to reliably simulate the EWA map for arbitrarily long time, since for a sufficiently large value of α the machine precision limits are never reached. The intuition for this property is simple: the performance of very successful or very unsuccessful actions is forgotten exponentially over time, so even a very small value of α prompts the players to choose unsuccessful actions with positive probability. The second advantage is that the EWA system is explicitly normalized every time step, making numerical errors unlikely.

EWA also has a computational disadvantage: because it uses exponential functions to map attractions into probabilities, if the value of the payoff sensitivity β is too large, the components of the mixed strategy vector may vary by too many orders of magnitude, and therefore overshoot the boundaries of the strategy space.

So care should be taken in choosing the values of α and β. This is the case also because of an additional feature of the EWA system: with large memory loss or small payoff sensitivity, the learning dynamics converge to the center of the strategy space. In the limit where $\beta = 0$ the players just choose uniformly at random between their possible actions, irrespective of the payoff matrix. In Galla and Farmer (2013) it was observed that for sufficiently large values of α/β a unique fixed point was always stable. Such a fixed point can be arbitrarily far from mixed strategy Nash equilibria, and so by changing their strategies the players can improve their payoff. We are not interested in this "trivial" attractor as we want to focus on the effect of the best reply structure of the payoff matrix on the learning dynamics. Therefore, we choose parameter values for α and β that prevent convergence to this fixed point.

[37] The value of 0.95 is smaller than the values of 0.98 and 0.99 used for reinforcement learning and fictitious play respectively. This is because finite memory and finite payoff sensitivity make it possible that a fixed point corresponding to a pure strategy Nash equilibrium is located at a small distance in the interior of the strategy space, as shown in Chapter 2.

A similar effect occurs for small δ. When the players only consider realized payoffs and do not give much weight to forgone payoffs they may get stuck in "lock-in" fixed points, in which they do not realize that by unilaterally changing action they could increase their payoffs (see Section 2.4.3). These fixed points are typically arbitrarily far from Nash equilibria. To avoid this situation, we assume that players consider forgone payoffs, and so choose $\delta = 1$.

A final important technical remark is that we rescale the payoff sensitivity β by $\sqrt{N}$ as the payoff matrix gets larger. The reason for this is that the expected payoffs $\sum_j \Pi^R_{ij} y_j$ and $\sum_i \Pi^C_{ji} x_i$ scale as $1/\sqrt{N}$. Indeed, focusing on the expected payoff of player Row, $\sum_j \Pi^R_{ij}$ scales as $\sqrt{N}$ due to the Central Limit Theorem (recall that the payoffs are generated randomly, see below for the precise rule), while the components y_j scale as $1/N$ due to the normalization constraint. So $\sum_j \Pi^R_{ij} y_j$ scales as $1/\sqrt{N}$. The same argument applies to the expected payoff of player Column. Now, note that β multiplies the expected payoff from Eqs. (1.29) and (1.32). Therefore, increasing the size of the payoff matrix has the same effect as decreasing β, until the attractor at the center of the strategy space becomes stable again. To prevent this from happening, we rescale β by $\sqrt{N}$, so that $\beta \sum_j \Pi^R_{ij} y_j$ and $\beta \sum_i \Pi^C_{ji} x_i$ do not scale with N.

In summary, for all simulations we choose $\alpha = 0.18$, $\beta = \sqrt{N}$, $\kappa = 1$ and $\delta = 1$, which ensure that the EWA dynamics stays within the strategy space, that it does not overshoot the space boundaries and that it does not reach the trivial attractor in the center of the strategy space or lock-in fixed points that have no relation to Nash equilibria.

3.A.5 Experience-Weighted Attraction with noise

As explained in Section 1.3.5, we study "online learning" as a model of stochastic learning. We consider the noisy version specifically of EWA because of its superior numerical properties (as compared to the other algorithms). Given that introducing noise makes identifying convergence more challenging, we choose the algorithm for which identifying convergence has been simplest.

Convergence criteria. As can be seen in Figure 3.19, the deterministic approximation of EWA and the noisy version are generally very similar. In the convergent example an action which is not the most commonly played one (i.e. the light green line) is selected from time to time, and this potentially pulls player Row away from equilibrium. What usually occurs instead is that the player returns to equilibrium after a short time.

We use the following convergence heuristic:

1. Only consider the last 20% time steps.

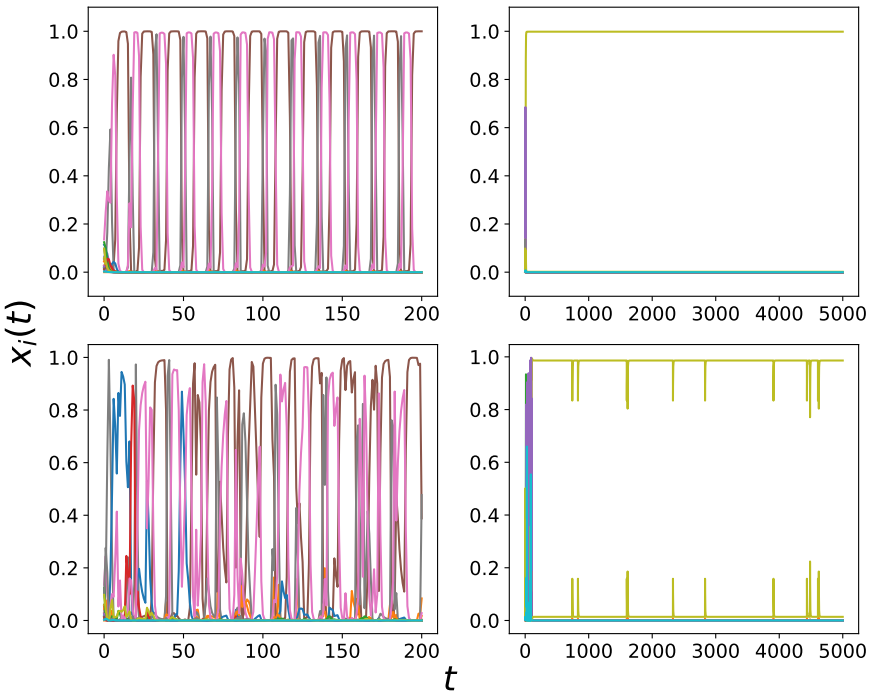


Figure 3.19: **Instances of simulation runs of EWA and EWA with noise** with $N = 20$. Each line is a component of the mixed strategy vector of Row (not all components are visible, as they overlap). Left panels: non-converging simulation run. Right panels: converging simulation run. Top panels: deterministic approximation of EWA. Bottom panels: EWA with noise.

2. Only keep the actions that have been played with a frequency larger than 0.05 (averaged over the time interval above).

3. Find the most common value of the probabilities, i.e. the fixed point.

4. Count the occurrences in which the probabilities are farther than 0.02 from the most common value.

5. If the occurrences are more than 10% of the considered time interval, identify the simulation run as non-convergent. Otherwise, identify it as convergent.

If the dynamics is highly chaotic and the number of actions is large, it is possible that point 2. above does not hold for any action, as all actions are only played occasionally. In this case we simply identify the simulation run as non-convergent.

Parameter values. Differently from the case of deterministic EWA, we need to consider a longer time interval for the dynamics to settle down to an attractor. We take 5000 iterations as a maximum, as for Bush-Mosteller dynamics and fictitious play. The values of the parameters are the same as for deterministic EWA.

3.A.6 Level-k learning

We consider the level-2 version of EWA, as explained in Section 1.3.6.

Convergence criteria. The dynamics is qualitatively very similar to EWA, so we use the same convergence criteria.

Parameter values. We also use the same parameter values. We assume that both Row and Column are level-2 players.

3.A.7 Initialization of the payoff matrices

In order to study generic payoff matrices, we sample the space of all possible payoff matrices by generating the payoff elements at random. Following Galla and Farmer (2013), at initialization we randomly generate N^2 pairs of payoffs (i.e., if Row plays i and Column plays j, a pair a, b implies that Row gets a, Column gets b), and we keep the payoff matrix fixed for the rest of the simulation (so the system described by the payoff matrix can be thought of as *quenched*). We consider an *ensemble* of payoff matrices constrained by the mean, variance and correlation of the pairs. The Maximum Entropy distribution that obeys these constraints is a bivariate Gaussian (Galla and Farmer, 2013), which we parametrize with zero mean, unit variance and correlation Γ. Therefore, $\Gamma < 0$ implies that the game is competitive (zero-sum in the extreme case where $\Gamma = -1$), while $\Gamma > 0$ encourages cooperation (see the main text). If $\Gamma = 0$ all best reply configurations are equiprobable because the payoffs are chosen *independently* at random, so we shall consider this as a benchmark case where we sample the space of all possible games with equal probability.

Part II

Endogenous fluctuations in macroeconomics

Chapter 4

Literature review

4.1 Overview

Business cycles have mostly been studied as the result of exogenous shocks. From a dynamical systems point of view, this means that deterministic trajectories in typical business cycle models would always converge to stable stationary states. Fluctuations occur because economies are recurrently hit by exogenous shocks, which pull them apart from stable states. A minority view, which is gaining new traction, holds instead that business cycles arise from strong forces that are endogenous to the economy and that, absent exogenous shocks, economies would nonetheless go through boom and bust periods. Mathematically, this means that business cycles are modeled at least in part by some form of non-linear dynamics—limit cycles or chaos.

We start this chapter by reviewing the literature that looks for statistical signatures of non-linear dynamics in economic time series data (Section 4.2). Since empirical evidence is controversial, we conclude that building theoretical business cycle models is a necessary complement to time series analysis. In Section 4.3 we review the literature on endogenous business cycle models. Because we identify a proliferation of forces that could cause business cycles, we introduce a framework based on recent work by Beaudry et al. (2019) that generalizes several of these forces. We next identify a key problem with business cycle models that explain fluctuations with exogenous shocks: these models have difficulties in accounting for the correlation of economic time series across sectors of the economy and across countries. In Section 4.4 we review the mechanisms that have been proposed to generate correlations, and augment our generalized framework by including the mechanisms that the empirical literature has identified. We then argue that endogenous business cycle models may be better suited than models based on exogenous shocks in accounting for sectoral and international correlations, because non-linear dynamics may align following the principles of synchronization theory. Accordingly, in Section 4.5 we introduce some

basic concepts of synchronization theory. Finally, in Section 4.6 we discuss the few works that have tried to apply synchronization theory to economics. We explain why these works have had limited impact, and why the contribution presented in this thesis may overcome these problems.

4.2 Empirical evidence for non-linear dynamics in economic time series

A widespread belief holds that there is little evidence for non-linear dynamics in economic time series. Given our definition of endogenous business cycles as being mathematically described by some form of non-linear dynamics, this lack of evidence seems rather problematic for the endogenous business cycle research agenda. We will argue in this section that empirical evidence is in fact mixed, not allowing either to support or to reject non-linear dynamics in the data. We will suggest that mathematical modeling of business cycles is a necessary complement to time series analysis.

Limit cycles. A first possible type of non-linear dynamics is limit cycles. In fact, business cycles were interpreted as periodic cycles until the 1950s. Of course time series were not perfectly periodic, yet the lack of clear periodicity was still justified by interpreting business cycles as a Fourier superposition of perfectly periodic cycles. As Slutzky (1937) put it:

> It is almost always possible to detect, even in the multitude of individual peculiarities of the phenomena, marks of certain approximate uniformities and regularities. The eye of the observer instinctively discovers on waves of a certain order other smaller waves, so that the idea of harmonic analysis, viz., that of the possibility of expressing the irregularities of the form and the spacing of the waves by means of the summation of regular sinusoidal fluctuations, presents itself to the mind almost spontaneously. If the results of the analysis happen sometimes not to be completely satisfactory, the discrepancies usually will be interpreted as casual deviations superposed on the regular waves. If the analyses of the first and of the second halves of a series give considerably divergent results (such as, for example, were found by Schuster while analyzing sunspot periodicity), it is, even then, possible to find the solution without giving up the basic concept.

As Slutzky observes, it is impossible, especially for relatively short time series, to reject the hypothesis of superposition of perfectly periodic cycles. One could always find justifi-

cations not to embrace the alternative hypothesis of random noise being at least partially responsible for the observed dynamics.

The main contribution in Slutzky (1937) was to show that fluctuations resembling the business cycle could also be obtained by averaging white noise in a way that corresponds to the modern concept of autoregressive process. Of course those fluctuations would not display peaks in the spectrum corresponding to the various harmonics of superposed perfectly periodic cycles. However, those fluctuations would have the property of recurrently oscillating below and above trend, as in actual business cycles. Together with the work of Frisch (1933), this lent credibility to the so-called Frisch-Slutzky paradigm of business cycles, at the basis of the exogenous business cycle view outlined in Section 4.1.

By the 1960s, the idea of business cycles as the superposition of regular cycles had fallen out of favor. This was further reinforced by the influential work of Granger (1966), which identified a "typical shape" in the spectrum of many economic variables. The typical shape is a spectrum density increasing monotonically from higher to lower frequencies, consistently with the spectrum of an autoregressive process of order one. Such spectra do not exhibit any peak, arguing against the existence of any periodicity in economic time series.[38] As Granger (1966) notes, peaks may be found in finite samples, but they would be spurious, i.e. not statistically significant.

Granger's finding would also rule out the possibility of stochastic limit cycles, i.e. dynamics defined from a deterministic limit cycle skeleton, augmented by some noise. Indeed, although stochastic limit cycles would not produce perfect harmonics in the spectrum, they would still produce significant peaks, which appeared to be ruled out by the typical shape.

Granger's hypothesis of a typical spectral density was backed by a number of macroeconomic time series ranging between the beginning of the 20th century and 1960. That period was characterized by two World Wars and an unprecedented global depression, therefore evidence for periodicity could not really be expected. As Granger himself noted, "It is not, of course, suggested that every economic time series produce such spectra." Yet, it appears that Granger's finding has gone undisputed.[39] Macroeconomics set out to explain fluctuations with models featuring stable steady states hit by recurrent exogenous shocks.

Very recently, Beaudry et al. (2019) re-examine this issue and show that humps in spectral density could in fact be an important property of some economic time series. Beaudry et al. (2019) claim that focusing on trending time series such as GDP and on

[38] The only exception is seasonality and its harmonics.

[39] A Google Scholar search of the papers that cite Granger (1966) shows that no paper which itself received at least 100 citations either confirms or disputes Granger's finding. The only exception is Englund et al. (1992), who find a concentration of spectral density between three and eight years in Swedish data, but do not test for its significance.

business cycles lasting between 2 and 8 years has led researchers to overlook evidence for periodicity. The authors focus on "cyclically sensitive variables"—variables, such as the unemployment rate, whose trend is small as compared to their cyclical fluctuations—and show that some of these display a statistically significant peak between 9 and 10 years in the US and other G7 countries. We will build on this evidence to propose a stochastic limit cycle model, and study its synchronization properties. However, a substantial part of our synchronization analysis also applies to chaotic dynamics, to which we now turn.

Chaotic dynamics. An alternative type of non-linear dynamics is chaos. We classify chaotic dynamics based on being low- or high-dimensional, although of course there is a continuum between the two. There exist several definitions of dimension of a chaotic attractor (Farmer et al., 1983), but for our purposes here we can just think of dimension as the number of degrees of freedom which are necessary to describe the dynamics. In the case of a limit cycle, the dimension is simply one, because dynamics follow a line. The dimension of a stochastic process, on the contrary, is equal to the dimension of the system, as dynamics are not constrained to any lower-dimensional manifold. The dimension of deterministic chaos is in between. Low-dimensional chaos retains some properties of limit cycles, for example it shows some periodicity. High-dimensional chaos, instead, is more similar to noise. Given the shortness of economic time series, it is practically impossible to disambiguate between high-dimensional chaos and noise.

Therefore, the literature has focused on testing for low-dimensional chaos, either deterministic or augmented with some noise. One of the most popular methods to test for chaos is computing a particular type of dimension known as the correlation dimension (Grassberger and Procaccia, 1983): A finite and greater than one value of the correlation dimension would indicate chaos. Another popular method is the calculation of Lyapunov exponents directly from time series data (Sano and Sawada, 1985; Eckmann et al., 1986): A positive Lyapunov exponent would indicate chaos. These two methods provide point estimates, so to assess statistical significance one needs to devise supplementary techniques. For example, Brock (1986) has provided a test for the correlation dimension, based on removing linear structure by calculating the correlation dimension on the residuals from the best fit autoregressive process.[40] If the data generating process was truly chaotic, one should still find the same correlation dimension as the one computed on the original time series. Similarly, Shintani and Linton (2004) have proposed a test for the statistical significance of positivity or negativity of Lyapunov exponents.

[40]This is not formally a statistical test as the distributional properties of the test statistic are unknown (Barnett and Serletis, 2000).

Applying these methods to economic time series to detect low-dimensional chaos, the empirical evidence is mixed.[41] Using the correlation dimension method, Brock (1986), Brock and Sayers (1988), Scheinkman and LeBaron (1989), Frank and Stengos (1988) and Frank et al. (1988) found little evidence for chaos in international economic and financial time series. However, Barnett and Chen (1988) and DeCoster and Mitchell (1991) did find chaos in some monetary aggregates. Using the Lyapunov exponent method, Shintani and Linton (2003) found statistically significant *negative* Lyapunov exponents in international real output series. However, Hommes and Manzan (2006), applying the same test to data generated by a low-dimensional chaotic model augmented with relatively strong noise, show that this could also produce negative Lyapunov exponents, implying that also low-dimensional chaos cannot be rejected.

Barnett and Serletis (2000) give a very good account of the controversies. Different tests have different null hypothesis and may have more power against different alternatives. The tests are very sensitive to dynamic noise. Time series are too short. More in general, these tests cannot find if chaos originates within the economy or from outside – e.g., in an agricultural economy it could just be that chaos originates within the weather system and then propagates into economic time series. The fact that the literature looking for evidence of chaos in economic time series subsided in the last 20 years suggests that complementary methods may be necessary to settle the controversies.

A possible way forward. Taking stock, it seems to us that rejecting evidence for non-linear dynamics in economic time series is too dismissive. It seems fairer to conclude that it is very difficult to provide conclusive evidence in favor or against non-linear dynamics by looking at the data alone. Since data alone cannot settle this issue, we need to build models that can feature stable steady states or non-linear dynamics according to the parameter values, and make distinct predictions in the two cases that may also concern other variables or other properties of the time series. If predictions from one of the two parameterizations were much more precise, this would constitute indirect evidence in favor or against non-linear dynamics.

4.3 Endogenous business cycle modeling

Endogenous business cycle modeling, based on obtaining fluctuations from non-linear dynamics, is at least as long a tradition as the modeling of business cycles based on stable

[41] The empirical evidence is stronger for non-linearity. Note that non-linearity does not necessarily imply non-linear dynamics: a stable system could be non-linear, and for example respond to positive and negative shocks in an asymmetric way. Because such systems are not of interest here, we neglect that part of the literature.

steady states recurrently hit by exogenous shocks. In this section, we give an account of endogenous business cycle models. We focus on explicit macroeconomic models, without considering models of part of the economy (cobweb models, financial models) that could be extended to macroeconomic settings. We only consider models for which the deterministic part of the dynamics is clearly separated from the stochastic component, so that it is clear if the model could yield substantial fluctuations driven by non-linear dynamics alone. This restriction rules out most complex macroeconomic Agent-Based Models (Dawid and Delli Gatti, 2018), for which the separation between the deterministic and stochastic components is not clear. Our survey will make clear how many complementary mechanisms have been proposed to generate endogenous business cycles. Given this proliferation of mechanisms, we will specify a general framework based on strategic complementarities (Beaudry et al., 2019), and construct a simple endogenous business cycle model that we will use in the following analysis.

4.3.1 A taxonomy of endogenous business cycle models

Models in which macroeconomic fluctuations resulted mainly from non-linear dynamics started to appear in the 1940s. Since then, hundreds, or even thousands, of models have been proposed, making a systematic review of the literature impossible. We will concentrate on a few papers only; interested readers can consult more comprehensive books and review articles, such as Lorenz (1993), Gandolfo (2009), and Boldrin and Woodford (1990). Endogenous business cycle models can be classified according to several dimensions.

In Table 4.1, models are grouped based on the economic force that causes endogenous fluctuations. Note that this is necessarily an approximation, as the instability of fixed points is caused by the interaction of many factors. The most common economic forces are:[42] (i) bounded rationality, corresponding to agents' expectations fluctuating about a fixed point, thereby producing fluctuations in economic activity; (ii) finance, corresponding to credit booms necessarily followed by credit crunches (Minsky, 1992); (iii) overinvestment, corresponding to firms investing more than necessary over a boom and then cutting back on investment causing a recession; (iv) technology, corresponding to clustering technological development in certain periods. Other economic forces that have been proposed to cause fluctuations are inventory behavior, monetary disturbances, overlapping generations effects, preference switch, search and matching, wage bargaining, and lags and delays of all sorts. As pointed out, we will restrain from focusing on one of these specific mechanisms by considering a general framework that can incorporate several of these forces.

[42]The narrative given below may not apply to all models in that class, but is given to build intuition.

Table 4.1: Some endogenous business cycle models

References	Economic aspects			Technical aspects		
	Force	Aggregate	Microfounded	Chaos	Continuous	Smooth
Bullard (1994)	Bounded rationality	YES	YES	NO	NO	YES
Schönhofer (1999)	Bounded rationality	YES	YES	YES	NO	YES
De Grauwe (2011)	Bounded rationality	YES	NO	NO	NO	YES
Bonart et al. (2014)	Bounded rationality	NO	YES	NO	NO	YES
Asano et al. (2019)	Bounded rationality	NO	NO	NO	NO	YES
Woodford (1988)	Finance	YES	YES	YES	NO	NO
Keen (1995)	Finance	YES	NO	YES	YES	YES
Azariadis and Smith (1998)	Finance	YES	YES	NO	NO	NO
Matsuyama (2007)	Finance	YES	YES	YES	NO	NO
Myerson (2012)	Finance	YES	YES	NO	NO	NO
Matsuyama (2013)	Finance	YES	YES	YES	NO	NO
Beaudry et al. (2019)	Finance	YES	YES	NO	NO	YES
Metzler (1941)	Inventories	YES	NO	NO	NO	YES
Sterman (1988)	Inventories	NO	NO	YES	YES	YES
Helbing et al. (2004)	Inventories	NO	NO	NO	YES	YES
Foley (1992)	Monetary	YES	YES	NO	YES	YES
Kaldor (1940)	Overinvestment	YES	NO	NO	YES	YES
Hicks (1950)	Overinvestment	YES	NO	NO	NO	YES
Goodwin (1951)	Overinvestment	YES	NO	NO	YES	NO
Scheinkman (1976)	Overinvestment	NO	YES	YES	NO	NO
Benhabib and Nishimura (1979)	Overinvestment	NO	YES	NO	YES	YES
Boldrin and Montrucchio (1986)	Overinvestment	NO	YES	YES	NO	NO
Lorenz (1987)	Overinvestment	NO	NO	YES	YES	YES
Benhabib and Day (1982)	Overlapping generations	YES	YES	YES	NO	NO
Grandmont (1985)	Overlapping generations	YES	YES	YES	NO	NO
Benhabib and Day (1981)	Preference switch	YES	NO	YES	NO	NO
Diamond and Fudenberg (1989)	Search and matching	YES	YES	NO	YES	YES
Judd (1985)	Technology	YES	YES	NO	NO	NO
Silverberg and Lehnert (1993)	Technology	NO	NO	YES	YES	YES
Matsuyama (1999)	Technology	YES	YES	NO	NO	NO
Horii (2012)	Technology	NO	YES	NO	YES	YES
Shleifer (1986)	Technology, Delay	NO	YES	NO	NO	NO
Day (1982)	Technology, Overinvestment, Bounded Rationality	YES	NO	YES	NO	NO
Goodwin (1967)	Wage bargaining	YES	NO	NO	YES	YES

A second classification is whether models are based on aggregate variables or not. In the third column of Table 4.1, we show which models are aggregate; for most disaggregated models, disaggregation is sectoral.[43] Disaggregation will be crucial in our model and for our analysis of synchronization.

A third classification is based on microfoundations. A broad definition of "microfoundation" is that the dynamical equations of the model are obtained by explicitly modeling the behavior of economic agents, rather than being directly assumed.[44] Since in economic theory the agents' behavior is most often based on optimization, and agents' expectations are always on average correct (*rational expectations*), the term microfoundations is usually narrowly used as a synonym of these assumptions. This is the meaning that we use in Table 4.1. Our model will not be explicitly microfounded, but we will discuss how it can be derived from microfoundations (in the broad or narrow sense) throughout.

Other classifications are based on more technical aspects. In the fifth column, we write whether the model admits chaos for some parameters or if it admits limit cycles only.[45] In the sixth column, we write whether the model is specified in continuous time or in discrete time. In the last column, we write whether the model produces smooth persistent dynamics or "sawtooth-like" dynamics, which "jumps" from one side of the attractor to another, with no persistence. Almost all continuous-time models produce smooth dynamics, while many discrete-time models produce sawtooth-like one. The empirical plausibility of these models is questionable, as economic time series clearly show persistence.[46]

For simplicity, the classification in Table 4.1 does not consider two interesting aspects. First, it may be useful to distinguish between business cycles and growth cycles, that is cycles on an otherwise increasing growth trend. The interaction between growth and fluctuations is sometimes important. For example, in the model proposed by Matsuyama (1999), the average growth rate is higher when the model allows endogenous fluctuations than when growth is monotonic. However, in most cases the growth and fluctuations dynamics are assumed independent, so this aspect is mostly superfluous. The second aspect that we did not consider is whether models are *indeterminate*. In the jargon of economic theory, indeterminacy (Benhabib and Farmer, 1999) is the property of some models that exhibit

[43] Exceptions are Sterman (1988) which, as all the literature on inventories, focuses on individual firms; and Asano et al. (2019), which consider disaggregate households.

[44] According to this definition, Agent-Based Models are microfounded (Dawid and Delli Gatti, 2018).

[45] We rely on what is reported by the authors only. Of course, it could be that the authors did not realize or write in their paper that their model could admit chaos in some parameter region that they did not explore or discuss.

[46] Proponents of these models justify the produced dynamics by identifying discrete time steps with long periodicities. For example, if the dynamics produces a cycle of period two, a time step could be interpreted as half business cycle periodicities. However, this procedure often leads to contradictions in terms of interpretation of the model of parameters (Boldrin and Woodford, 1990).

multiple equilibria (in some cases, a continuum of equilibria) to have paths that depend on the beliefs of the agents (i.e., are not determined from the *fundamentals* of the economy). Indeterminate equilibria can also be cycles. For example, in models such as Grandmont (1985), Azariadis and Smith (1998) and Diamond and Fudenberg (1989) there exist multiple cyclical equilibria, and which one is reached depends on the beliefs. Our model will be determinate (Beaudry et al., 2019), so we can safely ignore this aspect.

For its generality, and for its importance in the next section, we finally mention an important strand of research. This is the search for endogenous fluctuations in neoclassical growth models with microfoundations and rational expectations. These models are dynamic programming problems in which a representative agent maximizes utility over an infinite horizon. There is a parameter, the discount factor δ, that weighs future utility vs. current utility. The closer δ to one, the more the agent considers future utility. The closer δ to zero, the more the agent is "impatient", and only cares about current utility. A situation in which δ is close to one seems to be the ideal candidate for smooth growth paths: agents are infinitely foresighted and willing to give up current utility to avoid future variations in utility. And indeed, Scheinkman (1976) shows that for δ sufficiently close to one, endogenous fluctuations are not possible.[47] If this result were general, it would imply a tension between patience and instability, at least in models in which the expectations of the agents are rational. However, Boldrin and Woodford (1990) show that any deviation from the framework studied by Scheinkman (1976) makes endogenous fluctuations possible even if the discount factor δ is one. An example of such a deviation is given in the next section.

4.3.2 A generalized framework based on strategic complementarities

As should be clear from the discussion in the previous section, there exist a plethora of explanations for endogenous business cycles. To study synchronization of endogenous business cycles, it is not clear which specific model one might want to pick. It would be desirable to use a macroeconomic model that is as general as possible, so that it is clear that the synchronization results do not depend on a specific source for endogenous fluctuations.

Luckily, Beaudry et al. (2015) have recently proposed such a model.[48] They draw on research by Cooper and John (1988), who introduced the concept of strategic complementarities (Bulow et al., 1985) between agents in macroeconomics. Cooper and John (1988)

[47] The complementary result, proved by Boldrin and Montrucchio (1986), is that for any sufficiently small δ it is possible to specify a neoclassical growth model of this class that has endogenous fluctuations.

[48] Here we reference the initial working paper (Beaudry et al., 2015), rather than the final journal version (Beaudry et al., 2019), because it specifies the strategic complementarity framework in more detail.

themselves were dissatisfied by the plethora of models that were used to explain macroeconomic coordination failures, that is, situations in which the interaction among agents led to suboptimal outcomes. Some models at the time traced back coordination failures to increasing returns in firms' technologies, other models to imperfect information, others to search efforts of trading partners. Cooper and John (1988) showed that these situations could all be formalized within the unifying framework of strategic complementarities. Their framework was static, so Beaudry et al. (2015) make a dynamic version of it, and claim that it can encompass several of the causes for business cycles highlighted in Section 4.3.1.

Here we present a slightly modified version of the model in Beaudry et al. (2015), giving a more explicit role to heterogeneity across agents.[49] Consider N agents, indexed by $i \in \mathcal{I} = \{1, \dots, N\}$. Denote by x^i an accumulation variable representing a stock of agent i, and by y^i a decision variable representing a flow. We can think of y as level of production and of x as level of inventories; we can also think of y as level of investment or consumption of durables, and of x as capital or net worth, respectively.

The time evolution of x^i is simple. At each time step t, it depreciates by a factor δ, and increases by y_t^i. In formula, $x_{t+1}^i = (1-\delta)x_t^i + y_t^i$. Consider for example the interpretation of x as capital (machinery) and of y as investment. At every time step, capital depreciates, due to obsolescence and wear and tear in production. At the same time, old machinery is replaced by an amount of new machinery corresponding to investment y_t^i. The other interpretations for x and y lead to similar narratives. If x is the stock of inventories and y is production, new goods are added to the stock of inventories (and then potentially shipped to customers from there). If x is consumption of durable goods, such as cars or houses, and y is net worth, purchasing durable goods adds to one's personal wealth.

The evolution of y_t^i is more interesting, as it depends on the interactions among agents. Let $V^i\left(y_{t+1}^i, y_t^i, x_t^i, F\left(\overline{y}_t^i\right)\right)$ be a payoff function of agent i.[50] It first depends on the planned level of y, y_{t+1}^i, which can be set by the agent for time $t+1$ based on a number of variables at time t. These are the current level of y, y_t^i, the current level of x, x_t^i, and, most importantly, $\overline{y}_t^i$, capturing the interactions among the decision variables y of the agents. The effect of

[49] Beaudry et al. (2015) assume that agents' interactions are symmetric and consider the symmetric solution where all agents behave alike.

[50] A more general specification would be $V^i = \mathbb{E}_\tau^i \sum_{t=\tau}^{\infty} \beta^t V_t^i\left(y_{t+1}^i, y_t^i, x_t^i, F\left(\overline{y}_t^i\right)\right)$. In this more general specification, agents do not only care about the payoff at the current time τ, but also about future payoffs, with a discount factor $\beta \in [0,1]$. This makes agents *forward-looking*, in the sense that their expectations about future outcomes matter in the current period. (Here, $\mathbb{E}_\tau^i$ denotes the expectations of agent i at time τ.) A typical choice in macroeconomics to deal with this issue is to assume that expectations correspond to the outcomes of the model, at least on average (*rational expectations*). Mathematically, this corresponds to making the value of y at time t, y_t^i, depend on its future value y_{t+1}^i. Defining the dynamics consistently is a mathematical problem that led economists to invent new concepts that have no counterpart in the natural sciences, in which variables at time t only depend on variables at previous times. A recent contribution (Galizia, 2018) extended these methods to deal with limit cycles and chaos. In this thesis we ignore this aspect, but it could be possible to extend our results to the rational expectations framework.

this latter variable is mediated by a function $F\left(\overline{y}_t^i\right)$, which will be explained below. The interaction term $\overline{y}_t^i$ is defined as $\overline{y}_t^i = \epsilon_i^i y_t^i + \sum_{j=1}^N \epsilon_j^i y_t^j$, where $\epsilon_j^i \in [0,1]$, such that $\sum_j \epsilon_j^i = 1$. ϵ_j^i is the weight that the decision variable of agent j has in determining the value of $\overline{y}_t^i$. This of course includes the weight ϵ_i^i that the own decision variable of agent i has in setting $\overline{y}_t^i$. The values of ϵ_j^i, for all i and j, define an interaction network.

The function $F(\cdot)$, finally, is a non-linear function that determines the effect of interactions. Without loss of generality, let the payoff function $V(\cdot)$ depend positively on $F(\cdot)$. When $F'\left(\overline{y}_t^i\right) > 0$, there are *positive spillovers*, because an increase in the decision variables of the agents with whom agent i interacts raises the payoff of agent i. If such an increase also induces agent i to raise her own decision variable, $y_{t+1}^i > y_t^i$, one says that there are *strategic complementarities* between agent i and the agents with whom she interacts. Assuming that agent i wants to maximize her payoff (see below), she decides to raise her decision variable only if her marginal payoff increases with the decisions of the other agents. Mathematically, this means that the derivative of the payoff function $V(\cdot)$ with respect to the decision variable y_{t+1}^i is positive when the agents with whom agent i interacts increase their decision variables, so that $\overline{y}_t^i$ increases. In formula, $\partial^2 V^i\left(y_{t+1}^i, y_t^i, x_t^i, F\left(\overline{y}_t^i\right)\right)/\partial y_{t+1}^i \partial \overline{y}_t^i > 0$. If instead the function F is such that $F'\left(\overline{y}_t^i\right) < 0$, there are *negative spillovers*. If, in addition, due to an increase of the decision variables of the agents with whom agent i interacts, agent i reduces her decision variable, $y_{t+1}^i < y_t^i$, one talks about *strategic substitutability*. Mathematically, this condition corresponds to $\partial^2 V^i\left(y_{t+1}^i, y_t^i, x_t^i, F\left(\overline{y}_t^i\right)\right)/\partial y_{t+1}^i \partial \overline{y}_t^i < 0$. Depending on the value of y_t^i, our model will encompass both the complementarity and substitutability regimes. What will generate cyclical behavior in this model is the continuous switching between the two regimes.

Before we give more intuition on the dynamics, we complete the specification of the model. We first assume that the payoff $V^i\left(y_{t+1}^i, y_t^i, x_t^i, F(\cdot)\right)$ depends negatively on the current value of the accumulation variable x, so that $\partial V^i\left(y_{t+1}^i, y_t^i, x_t^i, F(\cdot)\right)/\partial x_t^i < 0$, and positively on the current value of the decision variable y, so that $\partial V^i\left(y_{t+1}^i, y_t^i, x_t^i, F(\cdot)\right)/\partial y_t^i > 0$. These assumptions reflect decreasing returns to accumulation, i.e. willingness to avoid excessive stocks of inventories, capital or net worth, and sluggishness in the adjustment of the decision variable, i.e. difficulty to quickly modify the level of production, investment, or consumption of durable goods. The first assumption will be instrumental for the switching between regimes of strategic complementarity and substitutability, the second assumption will be useful to introduce smooth dynamics. (As we discussed in Section 4.3.1, a problem of many endogenous business cycle models is that they generate a "sawtooth" dynamics that has no persistence, in stark contrast to empirical time series.)

Importantly for our analysis of synchronization, we let the dependence on current values of the accumulation and decision variables be heterogeneous across agents. This means that some agents dislike accumulation more than others, and some agents have more difficulty than others in adjusting their decision variables. This will generate heterogeneous frequencies of oscillation across agents.

We choose the following specification for the payoff function:

$$V^i\left(y^i_{t+1}, y^i_t, x^i_t, F\left(\overline{y}^i_t\right)\right) = -\left(y^i_{t+1}\right)^2/2 + y^i_{t+1}\left(\alpha^i_0 + \alpha^i_1 x^i_t + \alpha^i_2 y^i_t + F\left(\overline{y}^i_t\right)\right), \tag{4.1}$$

with $\alpha^i_1 < 0$ and $\alpha^i_2 \in [0,1]$, reflecting the assumptions discussed above. We will choose the parameters to restrict y^i_t to positive values, implying a positive dependence of $V^i(\cdot)$ on $F(\cdot)$, as assumed above. The dependence of $V^i(\cdot)$ on y^i_{t+1}, finally, should not be given any specific interpretation, as it is only instrumental to make y^i_{t+1} depend on the last term in parenthesis on the right hand side of Eq. (4.1). Indeed, one can assume that agent i chooses a value of y^i_{t+1} so to maximize her payoff $V^i(\cdot)$, and obtain a dynamical equation by setting $\partial V^i(\cdot)/\partial y^i_{t+1} = 0$.[51] This leads to $y^i_{t+1} = \alpha^i_0 + \alpha^i_1 x^i_t + \alpha^i_2 y^i_t + F\left(\overline{y}^i_t\right)$.[52]

We finally add two noise terms u^i_t and v_t in the evolution of the decision variables, corresponding to exogenous shocks. The first noise term is idiosyncratic to each agent, while the second term is common across agents. Both terms are AR(1) processes. This means that the idiosyncratic shock evolves as $u^i_{t+1} = \rho_u u^i_t + \iota_t$, where $\rho_u \in [0,1]$ is a persistence parameter and ι^u_t is white noise normally distributed with mean zero and standard deviation σ_u. The common shock evolves similarly, $v_{t+1} = \rho_v v_t + \iota^v_t$, where ι^v_t is the same white noise process, but with standard deviation σ_v.

Our model for endogenous business cycles is thus fully specified by the following equations, for all agents i:

$$\begin{array}{l} x^i_{t+1} = (1-\delta)x^i_t + y^i_t, \\ y^i_{t+1} = \alpha^i_0 + \alpha^i_1 x^i_t + \alpha^i_2 y^i_t + F\left(\overline{y}^i_t\right) + u^i_t + v_t, \\ \overline{y}^i_t = \epsilon^i_i y^i_t + \sum_{j=1}^N \epsilon^i_j y^j_t, \quad \epsilon^i_j \in [0,1], \quad \sum_j \epsilon^i_j = 1, \\ F(y) = \beta_0 + \beta_1 y + \beta_2 y^2 + \beta_3 y^3 + \beta_4 y^4. \end{array} \tag{4.2}$$

The last equation represents our arbitrary choice for $F(\cdot)$, that we parameterize as a generic quartic function (Beaudry et al., 2015). The choice of a quartic rather than a polynomial of any other order is due to technical reasons.

[51] This setting could be seen as a game, where each player i considers the strategic interactions with the other agents. For simplicity, we take the decision variables of the other agents as given, reducing the extent of strategic interactions.

[52] A minor difference with respect to Beaudry et al. (2015) is that y^i_{t+1} depends only on variables at t, while in Beaudry et al. (2015) it depended on both variables at t and (contemporaneous) variables at $t+1$. As there did not seem to be much difference in the dynamics, we chose this form as it was computationally simpler. Economically, it corresponds to assuming that the decision variables of the agents with whom agent i interacts have a lagged effect on her decision.

Although a mathematical analysis of the dynamics will be given in Section 5.2, here we build some intuition into why this dynamical system may generate cyclical dynamics. For simplicity, as we will do in Section 5.2, we focus on a single agent i whose dynamics is purely driven by her own decision variable, namely $\epsilon_i^i = 1$, $\epsilon_j^i = 0$, for all $j \neq i$. Assume that the parameters are chosen so that there is a unique steady state.[53] Further assume that the parameters of the function $F(\cdot)$ are chosen so that it increases at the unique steady state and below, while it decreases as y becomes large. This corresponds to strategic complementarity near the steady state and when the decision variable is small, and to strategic substitutability when the decision variable is large.

Starting from a situation below the steady state, the regime of strategic complementarity makes it optimal for agent i to increase her decision variable y^i, in a way that she overshoots the steady state. Explosive economic forces are at play. However, as agent i starts to overaccumulate, because the value of y^i more than offsets the depreciation $(1-\delta)x^i$, due to the negativity of the parameter α_1^i the rise of y^i slows down. Moreover, as y^i increases, the regime of strategic substitutability sets in, corresponding to implosive economic forces. Soon, y^i reverts and starts to decrease. At some point, the decrease in y^i will bring back to a regime of strategic complementarity, making agent i increase her decision variable again, and the process restarts.

Depending on the interpretation for the accumulation variable x, the decision variable y, and the function F, the narrative based on the alternance between strategic complementarity and substitutability could correspond to one of the economic forces causing endogenous business cycles listed in Section 4.3.1. Beaudry et al. (2019), for example, consider a microfounded model in which x is net worth, y is consumption of durable goods, and F models banks' willingness to give loans. In good times, lending is perceived safe, and so agents can borrow to consume more durable goods. This further strengthens the economic boom, making lending to be perceived even safer (strategic complementarity). When the economy slows down because of overaccumulation, lending instead starts to be perceived less safe, making banks cut back on credit (strategic substitutability). This behavior of the banks can easily trigger a recession, which lasts until agents have liquidated assets in excess, at which point the cycle starts again. The narrative in Beaudry et al. (2019) is similar to many of the narratives for finance-based endogenous business cycles listed in Table 4.1. However, the strategic complementarity-substitutability framework easily lends itself to other narratives, such as the ones based on overinvestment in the Keynesian tradition (Kaldor, 1940;

[53] This means that the strategic complementarities are not strong enough to generate indeterminacy (Benhabib and Farmer, 1999). See Galizia (2018) for more details about indeterminacy.

Hicks, 1950; Goodwin, 1951), and the ones based on temporal clustering of technological innovations (Judd, 1985; Shleifer, 1986; Matsuyama, 1999).

Note that so far we have been equating the dynamics of an agent i and the aggregate dynamics. While this was justified in Beaudry et al. (2015), who *assumed* that all agents behaved alike, it is not justified *a priori* in our setting. Indeed, our main contribution will be to show how agents may or may not synchronize their behavior and increase (or decrease) their decision variables at the same time. If they synchronized their behaviors, agents would produce aggregate business cycles but, if they did not, an aggregate business cycle would not *emerge*, because their behaviors would cancel out. In our model, synchronization will depend on the interaction network ϵ_j^i and on the nature of shocks. In the following section, we look at which economic forces may lead to synchronization and, accordingly, to what the interaction network should represent.

4.4 Channels for the synchronization of business cycles

In this section we look at which economic interactions may be responsible for the synchronization of endogenous business cycles. Following a literature review, we will conclude that trade is the least controversial source for synchronization across sectors and countries. Accordingly, we will consider a version of the model presented in Section 4.3.2 in which agents correspond to sector-country pairs, and their interaction pattern is the international input-output network. Our model will also account for common shocks as a synchronization mechanism.

4.4.1 Which economic interactions can lead to synchronization?

There is a large literature in economics about synchronization of economic activity. This literature uses the word "synchronization" casually, as a synonym for comovement, or correlation, of economic time series. Some of this literature is purely empirical, in the sense that it tries to explain the correlation coefficients among economic time series purely by econometric methods. Another part of this literature is more theoretical, in the sense that it builds models in which the propagation of exogenous shocks may reproduce the comovement across time series.

As we will explain in Section 4.5, our meaning for the word synchronization will be different, as we will take synchronization to mean "alignment of non-linear dynamics". Therefore, in this section we will use the words comovement and correlation to refer to the literature. The mechanisms that have been proposed to explain comovement will be our starting point for a theory of synchronization.

Although an agent in the model presented in Section 4.3.2 could be an individual firm or household, in this thesis we will consider some level of aggregation. In particular, we will consider agents to be economic sectors (industries) across countries, the basic unit of analysis being a sector-country pair.[54] Therefore, we will look for the literature on what may cause the comovement of sectors and countries.

Both the sectoral and international literatures are large. A common trait is that there is substantial disagreement in the empirical literature over the channels that may cause comovement, and theoretical models account only for a relatively small fraction of the empirical comovement.

Comovement across sectors. The comovement across sectors has always been considered a fundamental feature of the business cycle. In their pioneering definition of the business cycle, Burns and Mitchell (1946) wrote "a cycle consists of expansions occurring at about the same time *in many economic activities*, followed by similarly general recessions" (emphasis added). Empirically, all studies agree that the degree of comovement across sectors is quite high for several macroeconomic time series.

The first approach to explain this empirical regularity was to assume it, by tracing back aggregate fluctuations to aggregate shocks (Lucas, 1977). The argument was that since all sectors would be subject to these shocks, comovement would be high. Possible aggregate shocks were monetary policy, set by the central bank, or fiscal policy, determined by the government. Since some recessions were clearly not triggered by such aggregate shocks, researchers started to look for alternative explanations. One of the first papers to do so was Long and Plosser (1983), proposing a model in which shocks were hitting individual sectors. Such an assumption lead to more easily identifiable shocks; for example, political disturbances that would lead to an increase of the price of oil were clear sectoral shocks. In Long and Plosser (1983)'s model, sectoral shocks would diffuse through input-output linkages—requirements of a sector's goods, say oil, as intermediate inputs for the output of another sector—, leading to high comovement of output across sectors. Starting from this paper, some strands of the literature branched out that looked at the problem from different angles. Here we mainly focus on two strands of this literature, and briefly discuss two more.

The first strand investigates whether input-output linkages are indeed the main channel for sectoral synchronization, letting aside aggregate shocks. For example, Hornstein and Praschnik (1997) do find support for the input-output channel. As an alternative channel, Cooper and Haltiwanger (1990) and Murphy et al. (1989) propose consumption good

[54] An alternative choice could be region(state)-country pairs, but the relatively higher amount of data on sectors made us choose sectors. Considering regions would be an interesting extension for future work.

demand complementarities. These models assume that sectors produce final goods that can be consumed by households, instead of intermediate inputs for the production of other goods (as in Long and Plosser 1983). Because consumers are also workers, if things go well within a sector and workers in that sector get a pay increase, they will also be able to spend on consumption goods from other sectors, leading to positive comovement. Some other channels that have been proposed are home production (Benhabib et al., 1991) and intratemporal adjustment costs (Huffman and Wynne, 1999), with no general consensus on a single channel.

A second strand of the literature studies the emergence of aggregate shocks from sectoral shocks. A law of large numbers argument would hold that shocks hitting many small sectors would cancel out in the aggregate (Dupor, 1999). However, a large body of work shows that this argument can be flawed when a few sectors are suppliers to many sectors (Horvath, 1998, 2000; Acemoglu et al., 2012; Baqaee and Farhi, 2019). This literature focuses on the emergence of aggregate fluctuations, and so only indirectly on comovement. A partial exception is Carvalho (2010), which shows that the same mechanism that avoids sectoral shocks to cancel out in the aggregate can only explain about 30% of the comovement in the data.

There are two other strands of literature that are still worth mentioning. One strand attempts to fix a standard macroeconomic model—the Real Business Cycle (RBC) model—so that it could reproduce high sectoral comovement not only in output but also in employment. This literature is surveyed in Christiano and Fitzgerald (1998), which also contains a comprehensive analysis of the sectoral comovement in US hours worked. The second strand of literature focuses on factor models to study comovement (Forni and Reichlin, 1998; Foerster et al., 2011). These models find one or two factors that are common to all sectors, combined with many sectoral shocks. While these common factors could be interpreted as aggregate shocks (Lucas, 1977), they could also be interpreted as a common endogenous business cycle that synchronizes across sectors.

In sum, while the strong comovement across sectors for many economic variables is an uncontroversial empirical regularity, explaining such a comovement theoretically has lead to substantial controversies. From the literature surveyed here, we will take aggregate shocks, input-output linkages and final demand complementarities as possible channels for the synchronization of endogenous business cycles across sectors.

Comovement across countries. Even more controversial is the literature that attempts to explain the comovement across countries. Several papers proposed different channels, finding support for one or another depending on the method used and the data analyzed

(De Haan et al., 2008). Moreover, most international business cycle models in which countries are connected by trade links have been unable to account for more than a tiny fraction of the empirical degree of comovement. This problem was named by Kose and Yi (2006) the "trade-comovement puzzle". As we will discuss, taking a synchronization approach can contribute solving that puzzle.

There are five main channels that have been proposed to explain comovement across countries. The first channel is trade. Theoretically, this channel could increase or decrease comovement. It would increase comovement to the extent that it would propagate positive or negative shocks, but it would decrease correlations if it lead to specialization across countries (Frankel and Rose, 1998). The second channel is convergence in fiscal and monetary policy, which is particularly relevant in case of monetary unions. The third channel is similarity in industrial structure: if countries have very different industrial structures, they would respond to sector-specific shocks in different ways. The fourth channel is finance. From a theoretical point of view, like trade, finance could increase or decrease comovement. A positive correlation across countries would arise to the extent that financial integration makes financial institutions in several countries experience troubles at the same time, leading to contemporaneous credit crunches. Yet, high financial integration could also lead to negative correlation, because capitals could flow to countries where things are going better, exacerbating cyclical differences. Finally, a fifth channel is multinational corporate control. Multinationals could easily propagate shocks either through the transmission of know-how (technology shocks) or to deteriorating financial conditions of the parent/affiliate company (financial shocks).

The literature that explores the role of these shocks purely empirically, by using econometric methods to detect the effect of some channel on comovement, is summarized in Table 4.2. All studies used different methods and different datasets, so that it is not clear what one should attribute different results to. The least controversial channel for international comovement is trade: all studies find a statistically positive effect of trade on comovement.[55] Common fiscal and monetary policy and similar industrial structure are in some cases insignificant, but in other cases do have a positive effect. Finance is more controversial. While Imbs (2006) finds positive effects of financial integration on international comovement, Kalemli-Ozcan et al. (2013), by using a highly disaggregated dataset, conclude that, all else equal, financial integration leads to less comovement. Finally, the presence of multinational firms across countries always has a positive, and substantial, effect on comovement.

[55] The magnitude of this effect is more debated. While Frankel and Rose (1998) and Baxter and Kouparitsas (2005) find a very strong effect, Imbs (2004) and Inklaar et al. (2008) find a much smaller effect.

Table 4.2: Channels for international business cycle comovement.

Reference	Countries	Years	Method	Trade	Fiscal & Mon Policy	Similar Structure	Finance	Multinationals
Frankel and Rose (1998)	OECD	1959-1993	IV	+				
Clark and Van Wincoop (2001)	US & EU countries and regions	1963-1997	OLS-IV	+	\			
Kalemli-Ozcan et al. (2001)	OECD	1970-2000	OLS-IV					
Imbs (2004)	24 countries	1980-2002	Simultaneous equations	+		+	+	
Heathcote and Perri (2004)	US vs world business cycle	1972-2000	Summary statistics				–	
Desai and Foley (2004)	US vs world business cycle	1982-1999	OLS					+
Baxter and Kouparitsas (2005)	100 countries	1970-1995	Robust analysis	+	\	\		
Kose and Yi (2006)	OECD	1970-2000	IV	+				
Imbs (2006)	41 countries	1960-2000	Simultaneous equations				+	
Calderon et al. (2007)	147 countries	1960-1999	OLS-IV	+		+		
Inklaar et al. (2008)	OECD	1970-2003	Multivariate model	+	+	+		
Hsu et al. (2011)	OECD	1988-2002	Simultaneous equations	+	+	\		+
Kalemli-Ozcan et al. (2013)	OECD	1978-2006	OLS with many controls				–	
Kleinert et al. (2015)	France regions vs world countries	1993-2007	OLS with many controls	+		+		+
Cravino and Levchenko (2017)	34 countries	2004-2012	OLS with many controls					+
Di Giovanni et al. (2018)	France vs world countries	1993-2007	OLS with many controls	+				+

Notes: Countries column: When we indicate "XX countries", it means that the authors used a combination of advanced (OECD) and developing countries. Years column: The dates reported here are only approximate, because some papers use multiple datasets with different starting and final dates. Method column: In addition to using different methods for finding the effect of different channels on comovement, papers also differ in the way they process raw data. In this table we ignore the latter aspect. Channels columns: "+" stands for "positive effect"; "–" stands for "negative effect"; "\" stands for "insignificant effect"; no symbol means that the authors ignored that channel.

While the literature surveyed so far is purely empirical, some papers have proposed theoretical models to account for the empirical correlations across countries. These models build on the International Real Business Cycle (IRBC) framework of Backus et al. (1992). They focus on trade as a main channel for comovement, because financial linkages would decrease comovement in these models. Kose and Yi (2006) show that the basic IRBC model falls short of explaining the empirical degree of comovement, explaining only about 10% of the correlation in the data. They name this problem "trade-comovement puzzle". Subsequent work (Arkolakis and Ramanarayanan, 2009; Johnson, 2014; Liao and Santacreu, 2015) proposed several solutions to the trade-comovement puzzle, without much success. The key idea of the paper that got closest to explaining the puzzle (Johnson, 2014) is to disaggregate international trade by sectors, and explicitly model the trade of intermediate goods. This was motivated by previous work (Di Giovanni and Levchenko, 2010; Ng, 2010) which showed that the highest levels of pairwise correlations were experienced by the sector-country pairs with highest intermediate good trade intensity.

The main reason for the trade-comovement puzzle is that international trade is a small fraction of the total amount of transactions that happen within national borders. It is then difficult to obtain sufficiently high degrees of correlation by the transmission of exogenous shocks. A possible solution could be to consider financial linkages (Devereux and Yetman, 2010), but one should recall from Table 4.2 that the effect of financial integration is ambiguous. A more promising solution seems to include multinational corporate control (Cravino and Levchenko, 2017; Di Giovanni et al., 2018). As we will explain more in detail in the next section, this thesis proposes an alternative solution to the trade-comovement puzzle. We attack this problem by challenging the idea that comovement arises only from the transmission of exogenous shocks, proposing instead that it arises at least in part from the synchronization of non-linear dynamics that represent endogenous business cycles.

In sum, empirical studies have shown that international trade, common fiscal and monetary policy, and similarity in industrial structure are relevant channels for business cycle comovement. The role of finance is more controversial. Empirical and theoretical work has also shown that a promising solution to the trade-comovement puzzle is modeling international trade at the sectoral level. Building on this evidence, we will propose a model of synchronization in which agents are sector-country pairs and interactions are mediated by the international input-output network, which we describe in the next section. Our model will also allow for country-specific and sector-specific shocks, allowing to capture common fiscal and monetary policy and similarity in industrial structure. As our model can easily be extended to individual firms, an interesting extension of this work could be to account

Table 4.3: Example input-output table comprising two sectors in two countries.

		Buyer Sector						
		Country 1		Country 2		Final Demand		
Seller Sector		Sector A	Sector B	Sector A	Sector B	Country 1	Country 2	Total Output
Country 1	Sector A	5	10	20	10	45	10	100
	Sector B	10	5	10	20	50	5	100
Country 2	Sector A	30	15	800	500	5	8650	10000
	Sector B	35	30	1000	1000	25	7910	10000
Value Added		20	40	8170	8470			
Total Outlays		100	100	10000	10000			

Note: This example is taken from Cerina et al. (2015).

for the synchronizing effect of multinational corporate control, which was recently shown to be important.

4.4.2 An input-output endogenous business cycle model

Table 4.3 is an example input-output table. It comprises two sectors, A and B, in two countries, 1 and 2, which in this stylized example represent the entire world. We will refer to sector-country pairs as 1A, 2B, etc. The first four lines and four columns of the table represent the value of intermediate good flows across sectors. For a given entry of the table, the row is the seller sector, while the column is the buyer sector. For example, 1A sells intermediate goods that are worth 10 units of currency to 1B, so that 1B could use them for further production. Trade could also happen within sectors: firms in 1A sell goods whose total value is 5 to other firms in 1A. Trade can also happen internationally. As is clear already from the intermediate good flows, country 2 is much larger than country 1, and so unsurprisingly it is both the main customer and supplier of firms in country 1. The fifth row of the table represents value added of production of good 1A, obtained by summing employee compensation, interest payment for capital goods, taxes, entrepreneur profits, etc. Summing across the columns one can compute total outlays, obtained by adding the cost of buying intermediate goods to value added. The fifth and sixth columns represent final demand from countries 1 and 2, respectively. Final demand comprises consumption goods that are purchased by households, investment goods that are bought by firms, and government purchases. For example, final demand for sector 1A goods is worth 45 units of currency. Summing across the rows one obtains the total output of a sector, calculated from adding intermediate good sales to final demand. As an accounting identity, total output always corresponds to total outlays.

While the example above was very stylized, it captures the essence of input-output interdependencies. Figure 4.1 shows the network representation of an input-output table for

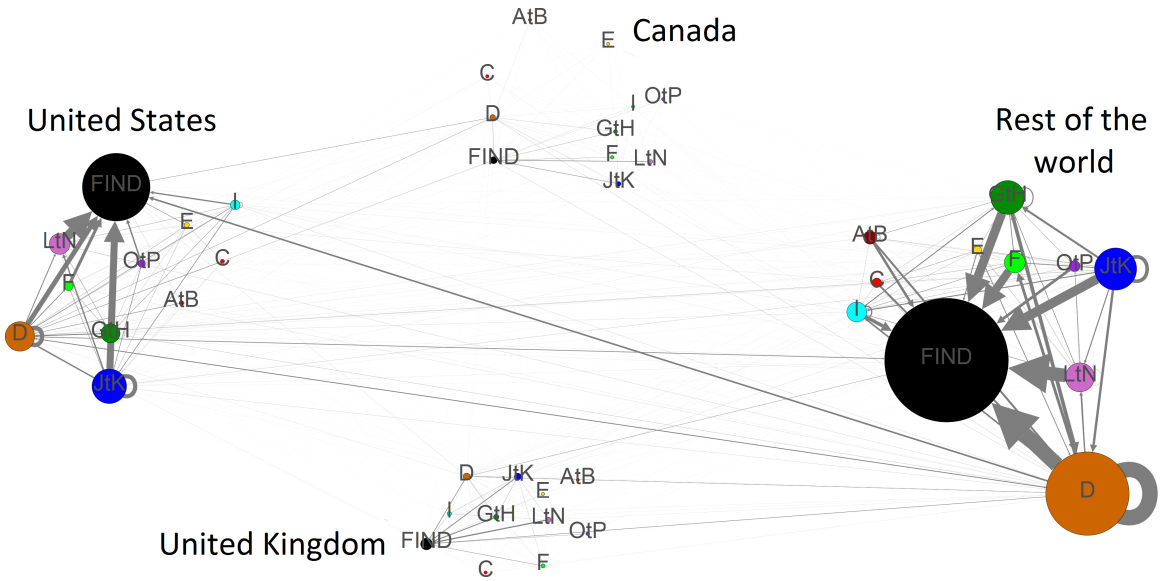


Figure 4.1: International input-output network in year 2000 showing three countries explicitly and the Rest of the World as an aggregate. Nodes are sector-country pairs and their size is proportional to their total value. Weighted directed links are value flows (goods and services flow in the same direction as the arrow) and their size is proportional to the value of these flows. Although value flows across sectors are bilateral, for visualization purposes arrows only indicate the largest direction. AtB: Agriculture; C: Mining; D: Manufacturing; E: Utilities; F: Construction; GtH: Trade; I: Transports; JtK: Finance and other business services; LtN: Government services; OtP: Other services; FIND: Final demand.

a real-world example.[56] This representation contains the same information as the equivalent table, except that we do not show value added. In network terminology, the total output of a sector-country pair is called its *out-strength*, while the total input of intermediate goods is its *in-strength*. Some stylized facts are immediately apparent. As in Table 4.3, some countries (the US) are much larger than others (UK and Canada). Some sectors, such as JtK - Business and LtN - Government, mostly supply final demand good and services, while sectors such as D - Manufacturing and GtH - Transport also supply intermediate good and services to other sectors. Finally, international trade seems to be much smaller as compared to trade within countries.

We now describe how we adapt the model described in Section 4.3.2 (Eq. 4.2) so that the interaction network is the international input-output network, discussing the narrative for endogenous fluctuations in this model at the end. In this specification, agents correspond to sector-country pairs ij, where i denotes sectors, $i \in \mathcal{I} = \{1, \ldots, I\}$, and j denotes countries $j \in \mathcal{J} = \{1, \ldots, J\}$. We also consider J final demand nodes, denoted by $0j$, i.e. $i = 0$ and $j \in \mathcal{J} = \{1, \ldots, J\}$, so that the total number of agents is $N = (I+1) \cdot J$. When considering firms in sector-country pairs, we can think of the decision variable y as production, and of the accumulation variable x as inventories. We think of final demand as being fully

[56] The data are taken from the World Input Output Database (Timmer et al., 2015).

based on consumption goods,[57] so that the final demand node in a given country is entirely "composed" of households within that country. Under such interpretation, the final demand interpretation of y can be consumption, while x can be interpreted as wealth.

One first challenge for the adaptation of the model is to account for the diverse size of sector-country pairs. To do so, we think of x_t^{ij} and y_t^{ij} as "intensive" or "capacity-utilization" variables fluctuating around the steady state. In parallel, we consider "extensive" variables $X_t^{ij} = O^{ij} x_t^{ij}$ and $Y_t^{ij} = O^{ij} y_t^{ij}$, where O^{ij} is a constant[58] representing total output of sector-country pair ij, for $i \in \{1, \ldots, I\}$, or total demand in country j, for $i = 0$. We choose parameters so that there exists a unique steady state denoted by $y_s^{ij} = 1$, so that, at the steady state, production of sector i in country j is indeed given by $Y_s^{ij} = O^{ij}$, and total final demand in country j is $Y_s^{0j} = O^{0j}$.

We now first consider the evolution of sector-country pair nodes, and after that we discuss final demand nodes. The evolution of intensive accumulation and decision variables x^{ij} and y^{ij} is analogous to the general specification in Eq. (4.2). In the following, we specify the interaction term. For $i > 0$, steady-state total output of sector-country pair ij is defined as $O^{ij} := \sum_{k=0}^{I} \sum_{l=1}^{J} W_{kl}^{ij}$, with W_{kl}^{ij} being the value of the flow of goods and services from sector-country ij to sector-country kl or final demand $0l$ in the steady state. The coefficients W_{kl}^{ij} can directly be read from international input-output tables (we assume the input-output network fixed over time). Total demand received by sector-country pair ij at time t is obtained by multiplying the steady-state value flows W_{kl}^{ij} by the capacity-utilization variables y_t^{kl} (including $kl = ij$),

$$\overline{Y}_t^{ij} = W_{ij}^{ij} y_t^{ij} + \sum_{\substack{k=0 \\ k \neq i}}^{I} \sum_{\substack{l=1 \\ l \neq j}}^{J} W_{kl}^{ij} y_t^{kl}. \tag{4.3}$$

We define the interaction term $\overline{y}_t^{ij}$ as intensive total demand received by sector ij, given by

$$\overline{y}_t^{ij} := \overline{Y}_t^{ij} / O^{ij} = w_{ij}^{ij} y_t^{ij} + \sum_{\substack{k=0 \\ k \neq i}}^{I} \sum_{\substack{l=1 \\ l \neq j}}^{J} w_{kl}^{ij} y_t^{kl}, \tag{4.4}$$

with $w_{kl}^{ij} := W_{kl}^{ij} / O^{ij}$. Given the definition of O^{ij}, $\sum_{k=0}^{I} \sum_{l=1}^{J} w_{kl}^{ij} = 1$; this means that the dynamics of sector-country pairs are identical to the ones of agents in Eq. (4.2), with the interaction structure specified by the input-output coefficients w_{kl}^{ij}.

[57] We do not consider investment as a component of final demand for two reasons. First, it is only around 20% of final demand. Since consumption has the largest share of final demand, ignoring investment makes sense as a simplifying assumption. Second, while there are abundant data on the intersectoral flow of intermediate goods, there are no systematic data of the intersectoral flow of capital goods.

[58] Our model does not consider growth.

An important point is that the input-output coefficients w_{kl}^{ij}, normalized by total *output*, are different from the technical coefficients of input-output analysis, which are rather normalized by total *input* (Miller and Blair, 2009). Technical coefficients are widespread in economics, and are for example used in the literature on aggregate fluctuations from idiosyncratic shocks (Horvath, 2000; Acemoglu et al., 2012; Baqaee and Farhi, 2019). Why do we use different coefficients? The reason is that models using technical coefficients are focused on *supply*. This makes sense, for example, in models of disruptions along value chains: if suppliers are unable to provide the required necessary intermediate good, firms have to cut back on production. Our model is instead *demand-driven*.[59] Using the same analogy, if customers demand less goods, firms cut back on production. While the debate about demand-driven and supply-driven models is unresolved in economics, we find the demand-driven specification more meaningful to model fluctuations based on strategic complementarities. This choice also gives a role to final demand, which would not matter in purely supply-driven models.

We now consider the evolution of final demand nodes. As for sector-country pairs, accumulation and decision variables evolve like in the general framework. The interaction term, however, is different. We assume that households, which in our specification fully determine final demand, work in the various sectors within their country, proportionally to the size of these sectors. Final demand then depends on the cyclical movements of sectors, weighted by their size. This is meant to capture the idea that sectors that are not doing well may lay off workers or reduce their wages, depressing final demand of those workers. Therefore,

$$\overline{y}_t^{0j} := \frac{1}{O^j} \sum_{i=1}^{I} O^{ij} y_t^{ij}, \quad O^j := \sum_{i=1}^{I} O^{ij}. \tag{4.5}$$

In network terms, final demand is just a node whose out-going edges are proportional to the out-strengths of the pointed nodes.

Finally, we allow for sector-country-specific shocks u_t^{ij}, sector-specific shocks v_t^i, and country-specific shocks w_t^j, all autoregressive processes of order one.

Summing up, the input-output specification of the endogenous business cycle model described in Section 4.3.2 is given by the following equations, for all $i \in \{0, \dots I\}$ and

[59] This also means that shocks only propagate upstream, in contrast with some models in the literature.

$j \in \{1, \ldots, J\}$:

$$
\begin{aligned}
&x_{t+1}^{ij} := X_{t+1}^{ij}/O^{ij} = (1-\delta)x_t^{ij} + y_t^{ij}, \\
&y_{t+1}^{ij} := Y_{t+1}^{ij}/O^{ij} = \alpha_0^{ij} + \alpha_1^{ij} x_t^{ij} + \alpha_2^{ij} y_t^{ij} + F\left(\overline{y}_t^{ij}\right) + u_t^{ij} + v_t^i + w_t^j, \\
&\overline{y}_t^{ij} := \overline{Y}_t^{ij}/O^{ij} = w_{ij}^{ij} y_t^{ij} + \textstyle\sum_{\substack{k=0 \\ k\neq i}}^{I} \sum_{\substack{l=1 \\ l\neq j}}^{J} w_{kl}^{ij} y_t^{kl}, \quad i \in \{1, \ldots, I\}, \\
&\overline{y}_t^{0j} := \tfrac{1}{O^j} \textstyle\sum_{i=1}^{I} O^{ij} y^{ij}, \quad O^j := \textstyle\sum_{i=1}^{I} O^{ij}, \\
&F(y) = \beta_0 + \beta_1 y + \beta_2 y^2 + \beta_3 y^3 + \beta_4 y^4.
\end{aligned}
\tag{4.6}
$$

This stylized model captures most of the channels that previous research found were useful to explain comovement (Section 4.4.1). In terms of sectoral comovement within the same country j, it captures (i) national shocks through the noise term w_t^j; (ii) trade of intermediate goods through the input-output network coefficients w_{kj}^{ij}, for $i > 0$; (iii) final demand linkages through the weight O^{ij}/O^j that each sector has in determining the evolution of final demand which, as is clear from Figure 4.1, is a central node. In terms of international comovement, the model captures (i) international trade of final and intermediate goods, disaggregated by sector, through the input-output coefficients w_{kl}^{ij}, for $j \neq l$; (ii) similarity in industrial structure, through sector-specific shocks v_t^i; (iii) common fiscal and monetary policy, by correlating national shocks w_t^j across a subset of countries. Given its richness and parsimony, this very simple model seems an ideal choice to mathematically illustrate the mechanisms for synchronization of endogenous business cycles across sectors and countries.[60]

The narrative for how this model could generate cyclical movements is similar to the one in Section 4.3.2. Strategic complementarity and substitutability regimes for households could correspond to banks' willingness to give loans, as before, with the additional element that financial conditions worsen most strongly when *large* sectors are performing badly, because this leads to a higher number of layoffs and pay cuts in the economy. To identify strategic complementarity and substitutability regimes for firms, we recall that the meaning of the interaction term w_{kl}^{ij}, $i > 0$, is trade of intermediate goods from sector-country pair ij to kl. In a strategic complementarity regime, ij wants to increase production when its customers kl also want to increase production, while the opposite is true in a strategic substitutability regime. This is very reasonable, and could be linked to the old Keynesian principle of the accelerator.

At the level of national economies, when the strategic complementarity regime sets in, firms increase their production and demand more intermediate goods to other firms, who increase their production themselves. At the same time, improving financial conditions make banks give more loans, which allow households to consume more, giving an additional

[60] We will discuss possible extensions of this model at the end of Chapter 5.

push to the production sector through final demand. This continues until households have overaccumulated and firms have overproduced, at which point the strategic substitutability regime sets in and a recession begins. After the liquidation of excess stocks of durable goods and inventories, the strategic complementarity regime sets in again, restarting the cycle. (Of course this process is not perfectly periodic, as sector-country pairs are recurrently hit by several types of shocks.) If trade links are sufficiently strong, this process could propagate internationally, making several countries' economies rise and fall at the same time.

For this narrative to make sense, the cycles across sectors within the same country and across countries need to become *synchronized*, i.e. sectors need to align their non-linear dynamics so that peaks and troughs are likely to happen at the same time, contrasting the de-synchronizing effect of idiosyncratic shocks. Although it is not obvious at all that synchronization could emerge, it is a really powerful mechanism to generate comovement. Indeed, it complements the main mechanism for comovement in exogenous business cycle models—similar propagation of shocks—with the alignment, and so the perfect correlation, of the deterministic component of the dynamics. In the next section, we review the basics of synchronization theory.

4.5 A crash course in synchronization theory

In a system made by two or more interacting components, synchronization is the adjustment of non-linear dynamics of the components due to their interaction. Synchronization was first discovered by Huygens in 1665. The Dutch scientist observed that two pendulum clocks hanging from opposite sides of the same wooden beam synchronized their motion in opposite swings. As he correctly noted, this phenomenon was due to the interaction of the two pendula that were imperceptibly moving the beam. The same phenomenon can be observed across a range of natural and social systems. One of the most fascinating manifestations of synchronization is the spontaneous contemporaneous flashing of hundreds of fireflies, which does not need any leader or cue from the environment to occur. Instead, it is sufficient that fireflies adjust their flashing rhythms in response to the flashes of the other insects. Many other examples can be found in Strogatz (2004).

Why is this phenomenon so ubiquitous across such diverse systems? The reason is that all these systems can be described in the common language of dynamical systems theory. It is a generic property of interacting components of a dynamical system to align their non-linear dynamics in a way that they operate in synchrony, provided some conditions are satisfied. The most common concept of synchronization, *phase synchronization*, applies to perfectly periodic oscillators that have heterogeneous frequencies and interact weakly. This concept refers to the possible state in which the frequencies of these oscillators become

identical—this is known as *frequency entrainment*—and the phases, although not identical, become *locked* to one another. The concept of phase synchronization can be extended to oscillators that are not perfectly periodic, such as dynamical systems following stochastic limit cycles or low-dimensional chaotic attractors. A stronger concept of synchronization entails not only the frequency and phase, but also the shape of the dynamics, for example their amplitude. In the most extreme case in which the non-linear dynamics of the units become identical, one talks about *complete synchronization*. This state is also possible for high-dimensional chaotic dynamics that have no evidence for periodicity.

In the following, we introduce some basic ideas of phase and complete synchronization. We will keep mathematical treatment to a minimum, as a more exhaustive analysis will be performed for dynamical systems (4.2) and (4.6) in Chapter 5. A standard reference for synchronization theory is Pikovsky et al. (2003).

4.5.1 Phase synchronization

Consider a trajectory $\boldsymbol{X}_t$ of an N-dimensional continuous-time dynamical system, following an arbitrary limit cycle of period T. Such a trajectory can first of all be characterized by its ordinary frequency $\nu = 1/T$ or angular frequency $\omega = 2\pi/T$. Letting t_0 be a time in which the trajectory is at an arbitrary starting point, it is also possible to characterize the trajectory through its phase $\theta(t) = 2\pi(t - t_0)/T \mod 2\pi$, where $\mod$ is the modulo operation. For constant phase differences, the trajectory generally travels different distances along the cycle, because velocity can be different at distinct phases.

Consider now two trajectories $\boldsymbol{X}_t^1$ and $\boldsymbol{X}_t^2$. They both follow a limit cycle in isolation. They can be characterized by their frequencies ω_1 and ω_2 and by their phases θ_1 and θ_2. Suppose now that the two trajectories become weakly mutually *coupled*, in the sense that the time evolution of $\boldsymbol{X}_t^1$ depends weakly on $\boldsymbol{X}_t^2$, and viceversa. An analytical description of the effect of coupling is impossible in the general case. However, it is possible to use a technique called *phase reduction theory* (Pikovsky et al., 2003; Nakao, 2016) to approximate the motion of each trajectory and the effect of coupling by a simple phase description of the dynamics. This approximation is only possible when coupling is weak and mostly the phase is affected (i.e., there are negligible effects on the amplitude).

Under further approximations, it is possible to describe the phase dynamics of the two trajectories by the following system of differential equations (Kuramoto, 1975):[61]

$$\begin{aligned} \dot{\theta}_1 &= \omega_1 + k_1 \sin(\theta_2 - \theta_1), \\ \dot{\theta}_2 &= \omega_2 + k_2 \sin(\theta_1 - \theta_2). \end{aligned} \tag{4.7}$$

[61] These equations are linear, which is seemingly in contrast with the statement that synchronization applies to non-linear dynamics only. However, one should consider that these equations are approximations of an oscillator that must be described by non-linear equations.

In the above equations, the coupling terms are $k_1 \sin(\theta_2 - \theta_1)$ and $k_2 \sin(\theta_1 - \theta_2)$. The parameters k_1 and k_2 are positive constants, while the sinusoidal terms tend to make the two phases and frequencies more similar to one another. Suppose, for example, that $\theta_2 > \theta_1$. Because the coupling term in the first equation is positive, this increases the effective frequency of the first trajectory, potentially making θ_1 "catch-up" with θ_2. It is instructive to investigate under which conditions this may happen. Denote the phase difference between the two trajectories by $\phi = \theta_1 - \theta_2$. By calculating $\dot{\phi} = \dot{\theta}_1 - \dot{\theta}_2$, replacing from Eqs. (4.7), it is possible to obtain a single differential equation

$$\dot{\phi} = \omega_1 - \omega_2 - (k_1 + k_2) \sin \phi. \tag{4.8}$$

The steady states $\phi^\star$ of the above equation are obtained by setting $\dot{\phi} = 0$, i.e. $\sin \phi^\star = (\omega_1 - \omega_2)/(k_1 + k_2)$. These steady states correspond to a situation in which the phase difference between the two trajectories becomes constant (*phase locking*), because their frequencies become identical (*frequency entrainment*).

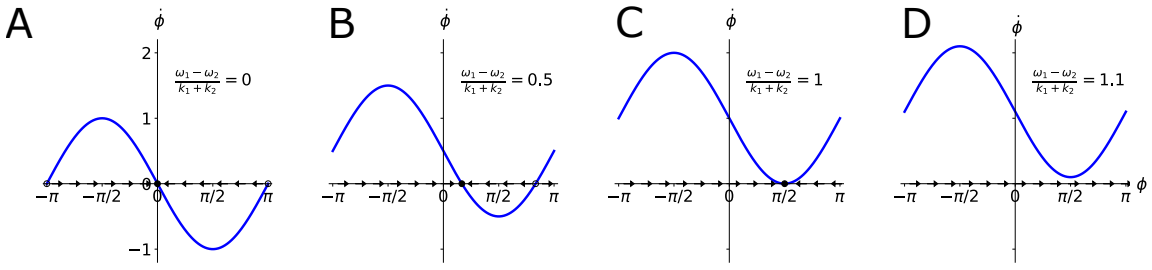


Figure 4.2: Existence and stability of steady states in the two-oscillators phase model as frequency spread becomes large compared to coupling (panels A to D). Full circles indicate stable steady states, while empty circles indicate unstable ones. Because of periodicity of all function, it is sufficient to study the interval $\phi \in [-\pi, \pi]$.

Steady states may or may not exist: they only exist if $-1 < \sin \phi^\star < 1$, i.e. if $|(\omega_1 - \omega_2)/(k_1 + k_2)| < 1$. This condition shows that for frequency entrainment to be possible, the frequency spread $(\omega_1 - \omega_2)$ must be small relative to the sum of the coupling coefficients $(k_1 + k_2)$. Also, steady states may or may not be stable: in Figure 4.2 we consider a few cases. In panel A, the uncoupled frequencies of the two trajectories are identical, $\omega_1 = \omega_2$ (or coupling is very strong), so that the only stable steady state is $\phi^\star = 0$. This suggests that, in the steady state, not only the frequencies, but also the phases, take the same values. The situation in panel B is more interesting, because the frequency spread is positive. There are two steady states, at $\phi^\star = \pi/6$ and $\phi^\star = 5\pi/6$, but only the first steady state is stable. The two trajectories now exhibit a constant *phase shift* $\theta_1 - \theta_2 = \pi/6$. Panels C and D show a situation in which coupling is not as strong as compared to the frequency spread. In panel C there is a unique stable fixed point at $\phi^\star = \pi/2$, while in panel D no stable fixed

points exist: the frequencies of the two trajectories remain different despite coupling, and the relative phases keep changing.

It is finally interesting to calculate the common frequency Ω on which the two trajectories converge if coupling is strong enough. This is

$$\Omega = \omega_1 \frac{k_2}{k_1 + k_2} + \omega_2 \frac{k_1}{k_1 + k_2}. \tag{4.9}$$

Somewhat unsurprisingly, the common frequency Ω is closer to ω_1 if the coupling coefficient k_2 is larger than k_1: if the second trajectory depends on the first more than the first trajectory depends on the second, it is intuitive that the common frequency should be closer to the frequency of the first trajectory.

Trajectories in our dynamical system (4.2) are generally strongly coupled, so that the approximations that are used to derive Eq. (4.7) do not hold (moreover, our dynamical system is a discrete-time map, rather than a continuous-time system of differential equations). We also consider many more trajectories than two, with a complex network as the interaction structure (Arenas et al., 2008; Rodrigues et al., 2016). We finally allow for noise, while the analysis so far has been based on purely deterministic systems. Nevertheless, we will see that many of the insights that one can obtain from the simple model studied in this section still apply.

4.5.2 Complete synchronization

Phase synchronization is a relatively weak synchronization concept, as it only involves the frequency and phase of trajectories. Periodic oscillations are also characterized by other variables, such as the amplitude. In the case of non-periodic trajectories, the usefulness of phase synchronization is less clear. If chaotic attractors are low-dimensional, it is still possible to extend phase synchronization analysis (Pikovsky et al., 2003). However, as chaos becomes more high-dimensional and any concept of periodicity disappears, there is no sense in which focusing on frequency and phase is useful. One could in fact think that synchronization in chaotic systems is impossible, because infinitesimally close trajectories exponentially diverge.

The concept of complete synchronization was in fact developed to study synchronization of chaos (Pecora and Carroll, 1990). Although the exponential divergence of nearby trajectories acts as desynchronizing factor, coupling works in the opposite direction. Complete synchronization theory quantifies the balance between synchronizing and desynchronizing factors by computing certain types of Lyapunov exponents. In its original formulation, complete synchronization was concerned with identical deterministic chaotic oscillators. However, the same insights hold for slightly different oscillators, or for oscillators hit by

(weak) noise. This theory was extensively used to assess the synchronizability of complex interaction networks (Pecora and Carroll, 1998; Barahona and Pecora, 2002; Arenas et al., 2008). In Section 5.4.1 we will build on complete synchronization to develop a theory that accounts for synchronization of non-linear dynamics, noise propagation and interaction structure. We will explain complete synchronization technically at that point.

4.6 Previous work on phase synchronization of business cycles

We are by no means the first who saw the appeal of synchronization theory to explain the origin and comovement of business cycles. As Krugman (1996) argues, in a popular science book:

> One of the luxuries of a format like this one is that I can include the kind of loose speculations that I could never write in a journal and that I can explain, as I am doing now, that I do not necessarily believe in the theory I am advancing. So here is a crazy idea about the global business cycle: it is an example of "phase locking."[...] Like the two back-to-back clocks that started ticking in unison, the two economies would not need to be very strongly linked to develop a synchronized cycle; a modest linkage would do as long as they were predisposed to have cycles in any case and had fairly similar natural periods.

In fact, it was not even Krugman himself who first considered applying synchronization theory to economics: the idea had probably been floating around since the first popularization of synchronization theory. Yet, for some reason, the few works that follow this idea have had very little impact in macroeconomics.

The main reason, we believe, is that these works had some weaknesses of one type or another. For example, Mosekilde et al. (1992) and subsequent works study mode locking between short-run business cycles and long-run cycles with 40-year periodicities. As discussed in Section 4.2, there is little empirical evidence for such long-run cycles. Selover and Jensen (1999) are the first, to our knowledge, to try and formalize Krugman's hypothesis of phase synchronization of international business cycles. However, they simply give economic interpretation to the simple two-oscillator model presented in Section 4.5.1, without developing an explicitly economically motivated model. More recently, Matsuyama et al. (2014) propose a micro-founded model of international synchronization of innovations. Yet,

the lack of ability of their model to generate smooth persistent dynamics questions its empirical applicability.[62] Finally, Gualdi et al. (2015) show how banks may synchronize and contemporaneously fail leading to economic crises, but their synchronization mechanism is very specific to the model they consider.

The model outlined in Section 4.4.2 addresses the issues with the previous papers at once. First, it has a clear economic motivation, based on strategic complementarities between sectors in various countries. Second, the fact that the model is not explicitly micro-founded highlights how the synchronization methods described in Section 4.5 can be generically applied to economic models. Third, it generates smooth persistent cycles which, when augmented with shocks, are reasonably similar to empirical economic time series. Fourth, and differently from all models presented above, it considers the full international input-output network, connecting sectors in different countries, at an arbitrary level of disaggregation. (The models above typically consider two countries or two sectors.) Fifth, we not only consider phase synchronization, but we also present a theory for complete synchronization, which works for the non-linear alignment of any dynamical system, even non-periodic ones. Moreover, our theory analytically assesses the impact of exogenous shocks. All the models above focus on phase synchronization, and at most check for robustness of their model to shocks.

[62] When the steady state is unstable, the model typically converges to a cycle of period two, although for some parameter values cycles of higher period are possible.

Chapter 5

Synchronization of endogenous business cycles

5.1 Overview

In this chapter, we provide a complete mathematical treatment of the synchronization properties of the general endogenous business cycle model described in Section 4.3.2 and of its input-output specification in Section 4.4.2. Our analysis confirms the intuitive narrative anticipated in Chapter 4, and provides new results.

In Section 5.2, we analyze dynamics of uncoupled agents. We show a transition from stability to instability through a Hopf bifurcation, in line with previous analysis by Beaudry et al. (2015). We then identify which parameters make the resulting limit cycle oscillate with higher or lower frequency. By assigning different values of these parameters to different agents, their oscillations become heterogeneous. Aggregating agents' uncoupled dynamics does not produce any aggregate business cycle.

In Section 5.3, we analyze what happens to heterogeneous frequencies when agents become coupled, for example via input-output demand linkages. We show that phase synchronization is possible—indeed, common—for sufficiently strong coupling or sufficiently weak frequency dispersion. When phase synchronization is achieved, aggregate fluctuations emerge from individual oscillations. While conditions on coupling and frequency dispersion are well known in the synchronization literature, our framework also yields new insights. For example, we show that different agents have different power in determining the common frequency of oscillation, depending on structural properties of the interaction network. This power has a clear economic interpretation. For example, the manufacturing sector has more power in setting the common frequency of oscillation than the sector of financial and business services, despite being smaller, because more sectors rely on demand of intermediate goods from the manufacturing sector than from the business sector.

The phase synchronization framework has some limitations, so in Section 5.4 we study complete synchronization of our business cycle model. We build on the so-called master stability approach (Pecora and Carroll, 1998) to introduce a theory that jointly accounts for synchronizing non-linear dynamics, exogenous shocks and interaction structure. This theory highlights the potential of endogenous business cycle models to generate relatively high comovement across economic time series. Indeed, comovement is the combination of perfect correlation of common non-linear dynamics and partial correlation due to similar propagation of exogenous shocks. The first non-linear channel is absent in models that only rely on exogenous shocks to generate comovement: this potentially explains the difficulty of these models to account for the empirical comovement across sectors and countries.

In this thesis we only give mathematical arguments for why synchronizing endogenous business cycle models instead have the potential to account for the empirical comovement. In future work, we plan to test this quantitatively. In Section 5.5, we list a number of extensions of the model that are likely to make it better match properties of the data.

5.2 Dynamics of uncoupled agents

We start the analysis by understanding the dynamics of uncoupled agents. For simplicity, we consider the general business cycle model introduced in Section 4.3.2, rather than the input-output specification (Section 4.4.2). We note from Eq. (4.2) that if we set $\epsilon_i^i = 1$, $\epsilon_j^i = 0$, for all j, it is $\overline{y}_t^i = y_t^i$, so dynamics of agent i only depends upon itself. Equivalently, we can focus on the symmetric solution in which the interaction network is symmetric and all agents behave alike, as in Beaudry et al. (2015).

We first look for the steady state (x_s^i, y_s^i). Following the discussion in Section 4.4.2, we constrain parameters so that $y_s^i = 1$ is the unique steady state. Noting that, at the steady state, the accumulation variable is $x_s^i = y_s^i/\delta$, for $y_s^i = 1$ to be a steady state it must hold $\alpha_0^i = 1 - \alpha_1^i/\delta - \alpha_2^i - F(1)$. In the following, we always select α_0^i so that this condition is satisfied. For the steady state to be unique, the slope of the line $-\alpha_0^i + y_t^i(1 - \alpha_1^i/\delta - \alpha_2^i)$ must be larger than the derivative of F at $y_s^i = 1$, namely

$$1 - \alpha_1^i/\delta - \alpha_2^i > F'(1). \tag{5.1}$$

The Jacobian of this system is

$$\mathcal{J} = \begin{pmatrix} 1-\delta & 1 \\ \alpha_1^i & \alpha_2^i + F'(1) \end{pmatrix}. \tag{5.2}$$

Stability of the steady state is completely characterized in terms of the trace $T^i = 1 - \delta + \alpha_2^i + F'(1)$ and determinant $D^i = (1-\delta)\left(\alpha_2^i + F'(1)\right) - \alpha_1^i$ of the Jacobian. Following the

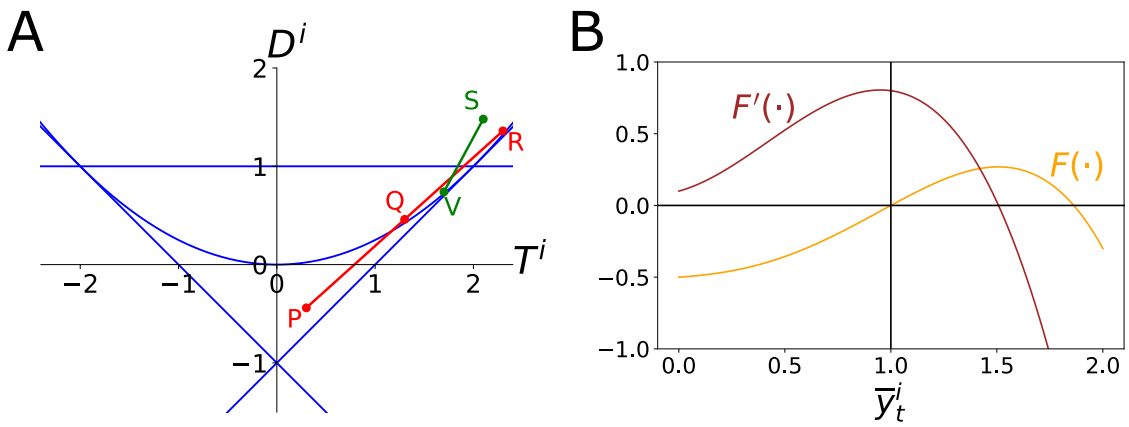


Figure 5.1: Stability of dynamics for uncoupled agents. (A) Diagram for stability of 2-dimensional maps. The blue lines correspond to $D^i = 1$, $D^i = T^i - 1$, $D^i = -T^i - 1$ and $D = (T^i)^2/4$. The red line corresponds to varying $F'(1)$ with the other parameters fixed, while the green line is obtained by varying α_1^i and α_2^i keeping the other parameters fixed. (B) Function $F(\cdot)$ and its first derivative $F'(\cdot)$ for $\overline{y}_t^i \in [0, 2]$.

standard conditions for 2-dimensional maps, stability obtains if $D^i < 1$, $D^i > T^i - 1$ and $D^i > -T^i - 1$. It is clear that dynamics for some agents may be stable and for other agents may be unstable, depending on whether these conditions are satisfied. If $D^i > (T^i)^2/4$, the eigenvalues of the Jacobian are complex, and so the system admits either damped oscillations or sustained cycles. If the line $D^i = 1$ is crossed above $D^i = (T^i)^2/4$, the system undergoes a Hopf (Neimark-Sacker) bifurcation.[63]

To build some intuition into the transition between stability and instability, in Figure 5.1A we plot the usual diagram for the stability of 2-dimensional maps. We first set the parameters $\alpha_2^i = 0.4$, $\alpha_1^i = -0.1$, $\delta = 0.1$, and vary $F'(1)$ in the interval $F'(1) \in [-1, 1]$. (The uniqueness condition (5.1) is satisfied for all these values of $F'(1)$.) Point P corresponds to $F'(1) = -1$. This corresponds to strategic substitutability, and indeed under this parameterization the steady state is a stable node. Point Q corresponds to $F'(1) = 0$, i.e. to a transition between strategic complementarity and substitutability. Under these parameters, the system has a stable focus. As the level of complementarities increases, the steady state loses stability through a Hopf bifurcation, up to point R, corresponding to $F'(1) = 1$. We then set the parameters of $F(\cdot)$ so that $F'(1) = 0.8$, we keep $\delta = 1$, and we vary $\alpha_1^i \in [-0.4, -0.02]$ and $\alpha_2^i \in [0.4, 0.]$. This produces the green line in Figure 5.1A; point S corresponds to $(\alpha_1^i, \alpha_2^i) = (-0.4, 0.4)$, point V to $(\alpha_1^i, \alpha_2^i) = (-0.02, 0)$. By looking at these bifurcation parameters, the transition from stability to instability occurs when there is a relatively strong dislike for accumulation and sufficiently high adjustment costs.

[63] Beaudry et al. (2015) show that the Hopf bifurcation is supercritical, i.e. the resulting limit cycle is attractive. While the dynamical system that we consider is slightly different (see footnote 52), numerical simulations indicate that the Hopf bifurcation is indeed supercritical.

The analysis so far has focused on local stability of the steady state. To build some intuition into global dynamics, in Figure 5.1B we plot the function $F(\cdot)$ and its first derivative. Here, and throughout, we follow Beaudry et al. (2015) and select $\beta_0 = -0.5$, $\beta_1 = 0.1$, $\beta_2 = 0.2$, $\beta_3 = 0.5$ and $\beta_4 = -0.3$. This choice guarantees that the function F is increasing at $\overline{y}_t^i = 1$ and decreasing when $\overline{y}_t^i$ becomes large, guaranteeing the existence a regime of strategic substitutability that prevents explosive dynamics.[64] Our analysis reproduces the narrative given in Section 4.3.2.

A last question concerning the dynamics of uncoupled agents is about how α_1^i and α_2^i determine the frequency of oscillation, in case the steady state is unique and unstable. For the following calculation, we drop the superscripts i. The eigenvalues are obtained from the trace T and determinant D of the Jacobian from the standard formula

$$\lambda_{1,2} = \frac{T}{2} \pm \sqrt{\left(\frac{T}{2}\right)^2 - D}. \tag{5.3}$$

In the region of the parameter space that we are interested about, the eigenvalues are complex, $D > T^2/4$. Denoting the negative argument of the square root on the right hand side in Eq. (5.3) by $M = -\left[(T/2)^2 - D\right]$, $M > 0$, the eigenvalues can also be written $\lambda_{1,2} = T/2 \pm i\sqrt{M}$, where i here denotes the imaginary unit. The frequency of oscillation is approximated by the argument ψ of the eigenvalues, $\psi = \arctan \operatorname{Im}\lambda / \operatorname{Re}\lambda = \arctan \sqrt{M}/(T/2)$. To calculate the effect of α_1 and α_2 on the frequency, we thus calculate the derivative of ψ with respect to these parameters. Starting from α_1,

$$\frac{\partial \psi}{\partial \alpha_1} = \frac{1}{1 + M/(T^2/4)} \cdot \frac{1}{T/2} \cdot \frac{1}{2\sqrt{M}} \cdot \frac{\partial M}{\partial \alpha_1}. \tag{5.4}$$

The first and third term on the right hand side are always positive. For the parameters that yield an unstable steady state the trace T is always positive, so also the second term is always positive. The last term is $\partial M/\partial \alpha_1 = -1$. Therefore, $\partial \psi/\partial \alpha_1 < 0$. Because we restricted $\alpha_1 < 0$, this means that values of α_1 that are larger in absolute value increase ψ. The economic interpretation is that a stronger dislike for accumulation increases the frequency, all else equal.

We now consider the effect of α_2:

$$\frac{\partial \psi}{\partial \alpha_2} = \frac{1}{1 + M/(T^2/4)} \cdot \frac{\frac{1}{2\sqrt{M}} \cdot \frac{\partial M}{\partial \alpha_2} \cdot \frac{T}{2} - \frac{\sqrt{M}}{2}}{T^2/4} = \frac{T\frac{\partial M}{\partial \alpha_2} - 2M}{(1 + M/(T^2/4))T^2\sqrt{M}}. \tag{5.5}$$

In the above equation, all terms in the denominator are always positive. We calculate

$$\frac{\partial M}{\partial \alpha_2} = \frac{1 - \delta - \alpha_2 - F'(1)}{2}. \tag{5.6}$$

[64] This choice of parameters also yields, conveniently, $F(1) = 0$.

For $\alpha_2 > 1 - \delta - F'(1)$, as is the case for the parameters we choose, this derivative is always negative, making the numerator on the right hand side on Eq. (5.5) negative (recall that $T > 0$ and $M > 0$). Therefore, $\partial\psi/\partial\alpha_2 < 0$. Since we restricted α_2 to take positive values, an increase of α_2 in absolute value leads to a decrease of ψ. In economic terms, all else equal, high adjustment costs lead to slower and smoother dynamics.

The difference in frequency due to different parameters α_1^i and α_2^i induces heterogeneous oscillations across agents. What can we expect in terms of business cycles?

5.3 Phase synchronization

We now couple agents and analyze how their frequencies of oscillation may—or may not—entrain. In Section 5.3.1 we study the generic business cycle model in Eq. (4.2) on an abstract interaction network. Some simplifying assumptions on the interactions make it easy to understand the effect of the topology of the network. In Section 5.3.2 we analyze the input-output model (Eq. 4.6) on the empirical input-output network, allowing to derive interesting insights on the sectors and countries that have most power in determining the common frequency of oscillation. All results in this section are obtained by numerical simulations, because the analytical methods surveyed in Section 4.5.1 require strong simplifying assumptions.

5.3.1 Abstract interaction network

We consider a special interaction network for which $\epsilon_j^i := \epsilon/K_i$ if j is connected to i and $j \neq i$, where K_i is the number of nodes i is connected to (degree of i). From the normalization condition it follows that $\epsilon_i^i = 1 - \epsilon$. This parameterization makes it possible to focus on the topology of the network. For simplicity, we further assume that the network is undirected, that is, if i is connected to j, also j is connected to i. The parameter ϵ makes it possible to quantify how much each node depends on the others. The larger ϵ, the stronger the dependence on the others. If $\epsilon = 0$, the dynamics of each node is independent from the others, and corresponds to the one described in the previous Section 5.2. We study a setting with $N = 10$ agents.

We select parameters so that all agents of the uncoupled system fluctuate, and study the effect of coupling on the entrainment of the various frequencies of oscillation.[65] We fix

[65] It is also interesting to study what happens in terms of stability when coupling stable and unstable dynamics. We find that small values of coupling tend to make stable agents unstable, but, as coupling strength increases, dynamics of all agents tend to become stable. The topology of the network has interesting effects: for example, the central node of a star network has more power in making the system stable or unstable, depending on its own stability. We do not report these results here as our focus is on synchronization.

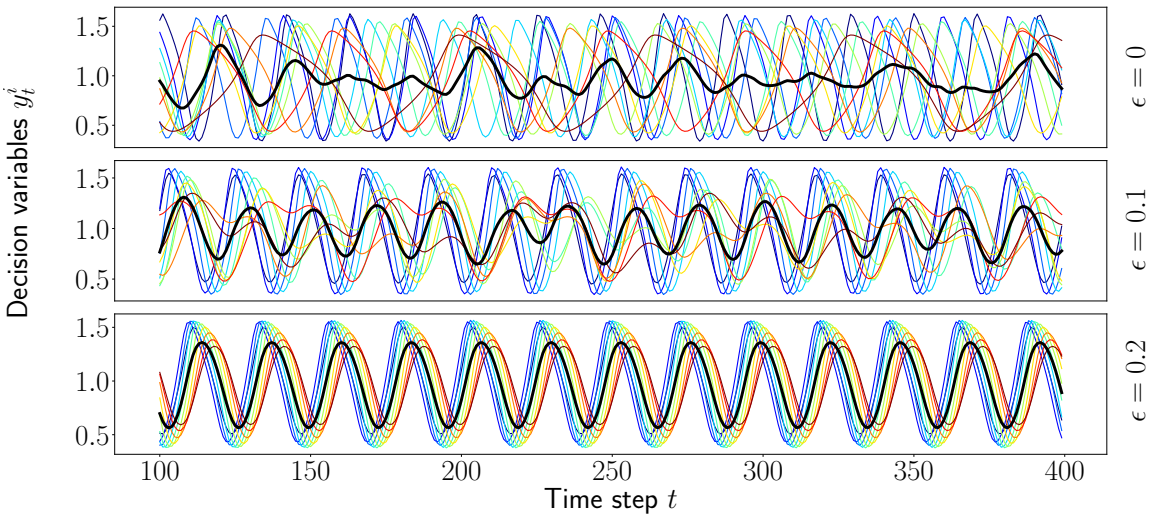


Figure 5.2: Time series generated by model (4.2) for a complete interaction network with $N = 10$ agents, for three values of the coupling coefficient ϵ. Colors from blue to red indicate time series that have higher to lower frequencies. Thick black lines are averages over the decision variables of the agents, which could be interpreted as aggregate business cycles.

$\delta = 0.1$, $\alpha_2^i = 0.4$ for all agents i, and vary α_1^i.[66] Agents whose dislike for accumulation α_1^i is stronger (in absolute value), when uncoupled, oscillate with higher frequency.

Deterministic cycles. We first analyze the case in which the noise terms in Eq. (4.2) are set to zero. We will then consider the effect of noise at the end of this section.

We give some intuition on the effect of coupling on dynamics in Figure 5.2. Here, we assign a distinct value of α_1^i to each agent by evenly spacing the interval $\alpha_1^i \in [-0.1, -0.02]$. In the top panel, corresponding to $\epsilon = 0$, the system is uncoupled. It is possible to see that the frequencies of oscillation are heterogeneous, with periods ranging from 20 time units to 66 time units. There is of course no indication of synchronization: the frequencies do not entrain, and the relative phases keep changing (there is no phase locking). Also, the shapes of the various oscillations are really different. Unsurprisingly, aggregating decision variables (thick black line) does not lead to any aggregate business cycle: aggregate fluctuations are much smaller than individual oscillations. In the middle panel, for $\epsilon = 0.1$, it is clearly possible to see that the frequencies have become similar, and some phases locked to one another. Yet, the shapes of some oscillations have remained quite different from the others, as can be seen from aggregate fluctuations still being quite smaller than individual oscillations. Finally, in the bottom panel the coupling strength is increased to $\epsilon = 0.2$. Now, not only the frequencies are the same and the phases are locked; also the shapes of oscillation have become identical, and phases are also very close across agents. The aggregate

[66] The results would be equivalent if we kept α_1^i fixed and varied α_i^2, to the extent that the frequencies produced were the same. We vary α_1^i because it produces a wider range of frequencies.

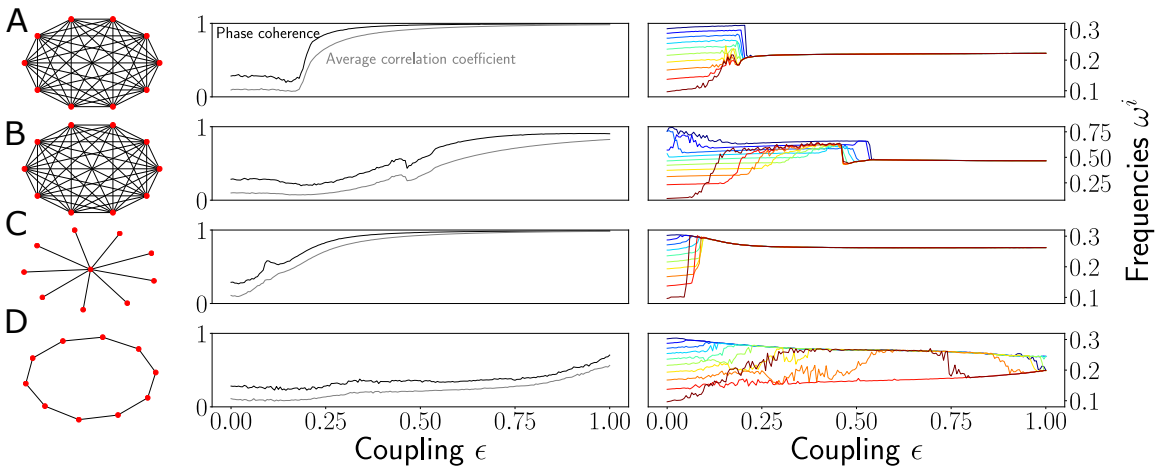


Figure 5.3: Phase synchronization in four cases A-D. Cases A and B are different because the frequency spread is higher in case B. Cases C and D have the same frequency spread as case A, but consider different interaction networks (star and chain, respectively). For all values of coupling $\epsilon \in [0, 1]$, we give: (i) phase coherence and average correlation coefficient; (ii) Frequencies ω^i of the agents. Highest (lowest) uncoupled frequencies are colored blue (red).

business cycle has a similar amplitude to individual oscillations. Further increasing ϵ make individual dynamics more and more alike.

While this example was useful to build intuition, we now analyze synchronization behavior more in detail. This is done in Figure 5.3. Each row represents a different combination of networks and frequency dispersion. In the first two rows, agents are connected in a complete network; in the third row, interactions occur through a star network; in the fourth row, the network is a chain. In the first, third and fourth row the values of α_1^i are evenly spaced in the interval $[-0.1, -0.02]$, leading to a frequency range $\omega^i \in [0.1, 0.3]$; in the second row, $\alpha_1^i \in [-0.4, -0.02]$, leading to $\omega^i \in [0.1, 0.8]$. This latter case corresponds to stronger frequency dispersion.

Panels in the first column represents two indicators that are commonly used to measure similarity between time series. The average correlation coefficient, typically used in the economics literature, is simply obtained by calculating the mean of all pairwise Pearson correlation coefficients between simulated time series. Phase coherence, popular in the synchronization community, measures instead similarity of phases. It is obtained by calculating the phase of each of the N oscillators at each time t, and then averaging across agents and across time. More precisely, focus on agent i. Letting t_l be the time step corresponding to the previous maximum of her decision variable y^i, and t_r be the time step of the following maximum, the phase ϕ^i at t is defined as $\phi_t^i = 2\pi(t - t_l)/(t_r - t_l)$. For example, if $t = 100$, and the closest maxima of the time series are at $t_l = 86$ and $t_r = 120$, the phase is $\phi_{t=100} = 2.58 = 2\pi \cdot 0.41$. Phase coherence at t is the modulus r_t of the complex number $r_t e^{I\Phi_t} = 1/N \sum_{i=1}^{N} e^{I\phi_t^i}$. ($I$ in the exponential is the imaginary unit.) When all oscillators

are at a similar point along the cycle (i.e. their phase is similar), r_t is close to one. If they are evenly spaced on the unit circle, r_t tends to zero. Phase coherence r is finally defined as an average over t. As we will see, the correlation coefficient and phase coherence give the same results.

Panels in the second column show how the different frequencies entrain, i.e. how they become similar to one another as the coupling strength ϵ increases. Each of the $N = 10$ lines represents the frequency of a given agent. Here and in what follows, we follow a color code that assigns blue to higher frequencies and red to lower ones.

Consider now the first row, case A. It is apparent that synchronization occurs abruptly. It is well-known in synchronization theory that this could happen, especially when the number of agents is large: it is an example of a *phase transition*. Both phase coherence and the average correlation coefficient increase abruptly at $\epsilon = 0.2$. Also the frequencies entrain at the same value of ϵ. Interestingly, the first frequency that adjusts to the others is the lowest frequency (largest periodicity), while the highest frequency is the last one to adjust. The frequency of the coupled system is close to the arithmetical average of the frequencies of the uncoupled agents.

The second row (case B) is analogous to the first row except that the parameters α_1^i are selected in the larger interval $[-0.4, -0.02]$, leading to stronger frequency dispersion. Unsurprisingly, synchronization is harder, and is obtained only for a much larger value of ϵ. As in the previous case, the first frequency that adjusts is the smallest one. Interestingly, there is little evidence for a phase transition, as the transition to synchronized behavior is smoother (but this could be due to finite size effects, given that we only have $N = 10$ agents).

In the third row (case C) the agents have the same parameters as in the first row ($\alpha_1^i \in [-0.1, -0.02]$), but the interaction network is a star. This situation is the one in which synchronization occurs first. This is due to the fact that all agents are connected to the central node of the star network, which acts as a "hub". Indeed, it is possible to see in the figure that when frequencies adjust, they first adjust to the frequency of the central node, as opposed to adjusting to the average frequency (cases A and B). These results are in line with known results about synchronization of "scale-free" networks (Arenas et al., 2008), which have many hubs that have a similar effect to the central node in the star network that we consider. As an arbitrary choice, we have assigned the central agent the highest frequency. Interestingly, we also find that the common frequency of oscillation—the one at which all agents converge is ϵ is high enough—is no longer the arithmetical average between the uncoupled frequencies, but is much closer to the uncoupled frequency of the central node. This is much less noticed in the literature, although it is fully consistent with

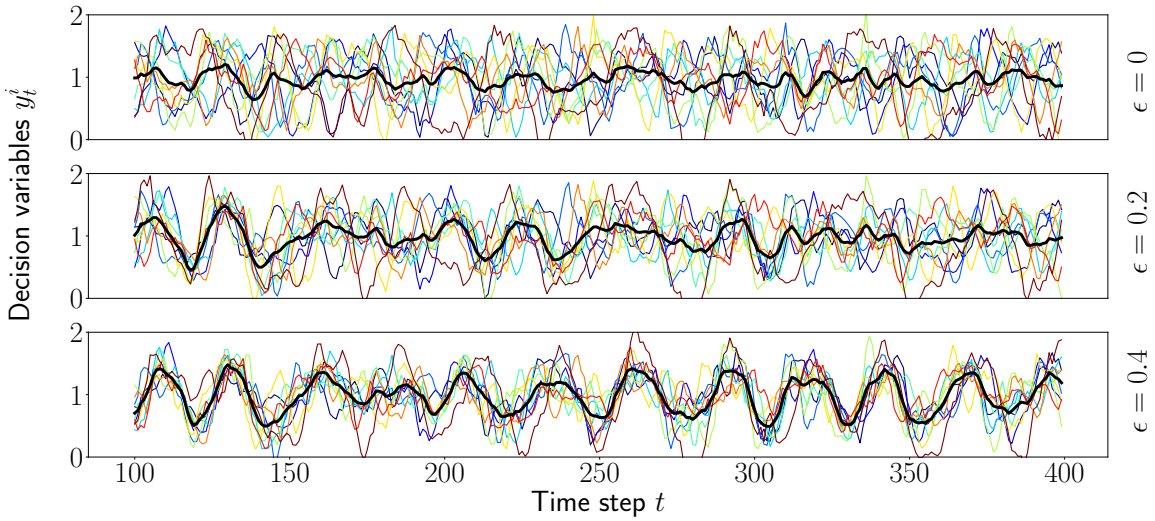


Figure 5.4: Time series generated by model (4.2) augmented by idiosyncratic noise u_t^i. As in Figure 5.2, we consider a complete interaction network with $N = 10$ agents, for three values of the coupling coefficient ϵ.

the simple Kuramoto example in Section 4.5.1 (the central node is weakly coupled to the leaf nodes, while the leaf nodes are strongly coupled to the central node).

Finally, the fourth row (case D) is a situation in which interactions are mediated by a chain network. This situation is the one in which synchronization is most difficult. The most likely reason is limited interactions due to local connectivity.[67]

Stochastic cycles. There are two noise terms in Eq. (4.2). Idiosyncratic noise u_t^i can be expected to act as a desynchronizing force, since, by definition, it affects each agent differently. It is a priori unclear, however, to which extent phase synchronization is possible under idiosyncratic noise. Eq. (4.2) also contain a common noise term v_t. It is well known in synchronization theory that common noise makes frequency entrainment possible even in the absence of coupling. The synchronization literature typically deals with noise by transforming phase equations into stochastic differential equations, and showing that the main results of synchronization theory are robust to noise (Pikovsky et al., 2003). Analytical results make strong simplifying assumptions on the type of noise and coupling. Here, we address synchronizability under noise by performing numerical simulations.

Figure 5.4 shows example time series, under relatively strong idiosyncratic noise, parameterized as an AR(1) process with persistence $\rho_u = 0.5$ and standard deviation of the (normal) innovation term $\sigma_u = 0.1$. As in Figure 5.2, the top panel shows dynamics of isolated agents ($\epsilon = 0$). Differently from the deterministic case, idiosyncratic shocks make

[67] The seminal paper on small-world networks by Watts and Strogatz (1998) was precisely conceived to enhance synchronizability of chain networks by creating "shortcuts" across the network.

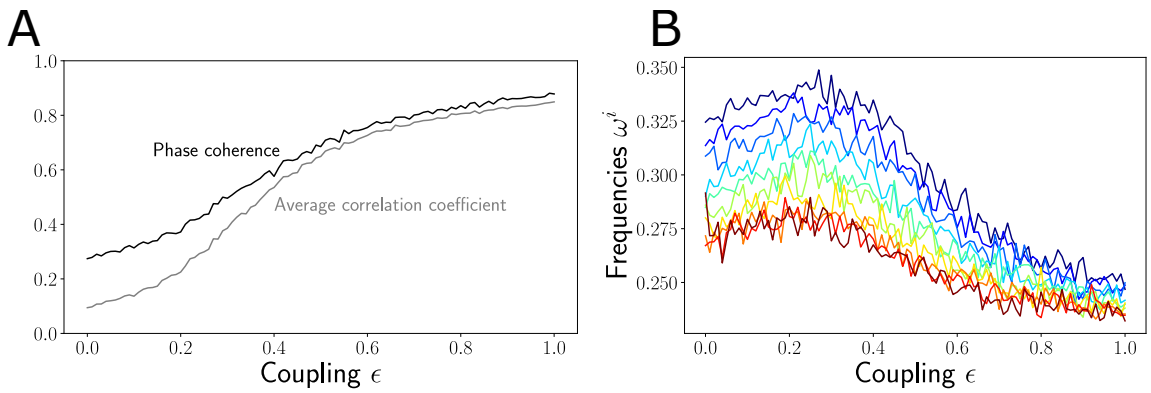


Figure 5.5: Phase synchronization under idiosyncratic noise u_t^i. Panels A and B can directly be compared to the equivalent deterministic scenario in the first row of Figure 5.3.

individual fluctuations resemble economic time series much more. Yet, similarly to the deterministic scenario, no aggregate business cycle emerges, as individual fluctuations cancel out. We then consider stronger coupling $\epsilon = 0.2$. This was more than enough to achieve synchronization in the deterministic system, but it is not enough here. We need to assume much stronger coupling, $\epsilon = 0.4$, to obtain a sufficiently strong level of synchronization. This also confirms, however, that phase synchronization is still possible under idiosyncratic shocks.

Further confirmation comes from the systematic investigation in Figure 5.5, in which we measure phase and frequency for increasing coupling ϵ.[68] The parameters are the same as under case A (first row) in Figure 5.3, except that we set the same level of idiosyncratic noise as above. Comparing this scenario to the scenario in Figure 5.5A, under all diagnostics synchronization is weaker. Here, phase coherence and the average correlation coefficient are at most around 0.8, while complete frequency entrainment is not obtained even for $\epsilon = 1$. Results are similar for the other interaction structures (star and chain network). This weaker level of synchronization matches the non-perfect comovement of empirical economic time series.

We now consider common noise v_t^i. To focus on the synchronizing effect of common noise, we exclude any other coupling by setting $\epsilon = 0$. We keep the correlation of the AR(1) process fixed to $\rho_v = 0.5$, while we vary the standard deviation of the innovation term. This

[68] Measuring phase and frequency is more difficult than in the deterministic scenario, because time series display many "local" maxima and minima. Suppose that a positive shock hits at a time step t at which dynamics of agent i is in a downward phase. The positive shock temporarily reverses the dynamics, but pretty soon the deterministic component of the dynamical system reinstates the downward phase. It would thus be misleading to identify time t as a turning point. To identify "global" maxima and minima we consider a window of observation t_w: any y_t^i that is larger than all points within the window of observation, $y_t^i > y_v^i$, for all $v \in [t - t_w, t + t_w]$, is identified as a global maximum. (The definition for global minima is equivalent.) In the figure we choose $t_w = 7$, which is anyway much smaller than the shortest period of the deterministic system (around 20 time units).

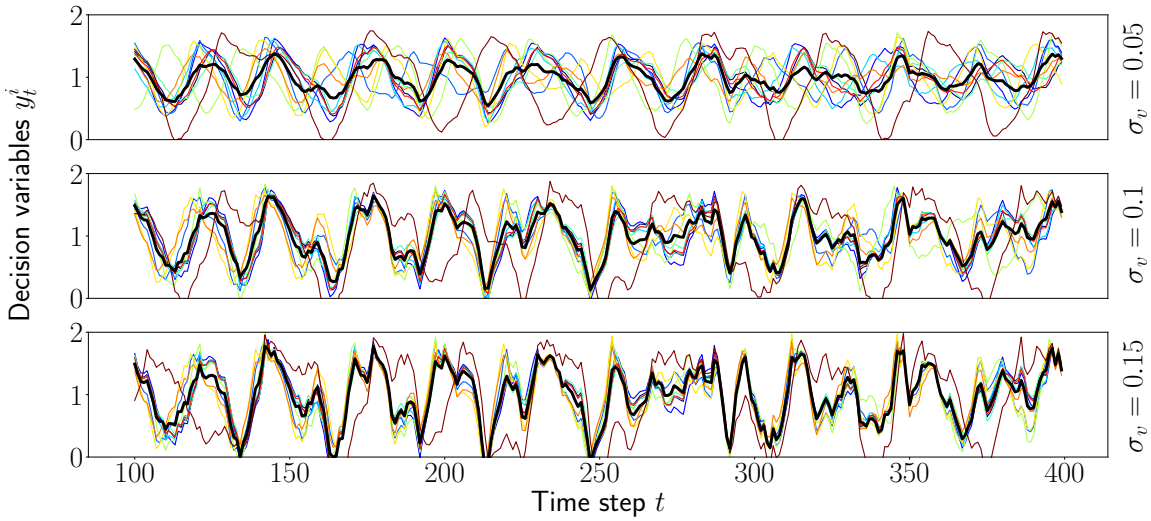


Figure 5.6: Time series generated by model (4.2) augmented by common noise v_t^i, with standard deviation σ_v. No further coupling exists: $\epsilon = 0$.

takes increasing values $\sigma_v = 0.05$, $\sigma_v = 0.10$ and $\sigma_v = 0.15$. In Figure 5.6 we show that stronger common noise leads to stronger synchronization, but also to weaker periodicity. Interestingly, while the first frequency that adjusts to the others under coupling is the lowest one (see Figure 5.3), under common noise the lowest frequency is the one that adjusts last. These results support common shocks (Lucas, 1977) as a possible synchronization channel, although the level of synchronization that is achieved under common noise is not as strong as under coupling.

5.3.2 Input-output business cycle model

We now study phase synchronization of the input-output model described in Section 4.4.2 (Eq. 4.6). Differently from the abstract interaction network that we considered in the previous section, the interaction coefficients w_{kl}^{ij} are highly heterogeneous, reflecting the fact that some nodes are "larger" than others. This property of the interaction network will lead to a number of interesting insights. We first focus on the US input-output network in isolation, by removing all the edges pointing to other countries, which in any case account for only about 5% of the value flows. We then consider the opposite case in which we aggregate all sectors within any given countries and study the international synchronization of business cycles. We finally discuss the complete case in which the basic unit of analysis is a sector-country pair. In this section we only consider deterministic dynamics, results for stochastic dynamics are similar.

Synchronization centrality in the US input-output network. Starting from the raw data, we calculate the interaction coefficients w_{kj}^{ij}, for $i, k \in [0, \ldots, I]$ and j corresponding to the United States. Following our discussion in Section 4.4.2, we obtain the interaction coefficients for sectors by dividing the value flows by the total output of each sector. We instead calculate the interaction coefficients for final demand by considering the output share of each sector. In network terms, final demand is a node with outgoing weighted edges to all other nodes, such that the weight of the edges is proportional to the total out-strength of the pointed nodes.

Table 5.1 shows the interaction matrix for 10 macro sectors and final demand (it corresponds to the network on the left in Figure 4.1). The first 10 rows in Table 5.1 are the shares of output that flow from the row to the column. For example, 0.58 in the first row means that 58% of "AtB - agriculture" output flows to manufacturing, and 0.18 in the same first row means that 18% of agriculture output is directly consumed ("FinD" - Final Demand). The 11-th FinD row simply shows the output shares of the various sectors. The two largest sectors are "JtK - Finance and other business services", with 28% of the output share, and "D - Manufacturing", with 21%.

Table 5.1: Input-Output US

Sector	AtB	C	D	E	F	GtH	I	JtK	LtN	OtP	FinD
AtB	0.22	0.0	0.58	0.0	0.0	0.01	0.0	0.0	0.01	0.0	0.18
C	0.0	0.04	0.49	0.25	0.04	0.0	0.01	0.01	0.02	0.0	0.12
D	0.01	0.01	0.35	0.0	0.06	0.05	0.03	0.03	0.07	0.01	0.39
E	0.02	0.02	0.21	0.0	0.01	0.1	0.02	0.12	0.1	0.02	0.38
F	0.0	0.0	0.01	0.01	0.0	0.01	0.01	0.05	0.04	0.0	0.87
GtH	0.0	0.0	0.1	0.0	0.04	0.04	0.01	0.05	0.04	0.01	0.7
I	0.01	0.0	0.12	0.03	0.03	0.09	0.14	0.11	0.08	0.03	0.36
JtK	0.0	0.01	0.08	0.01	0.02	0.07	0.03	0.25	0.08	0.03	0.42
LtN	0.0	0.0	0.01	0.0	0.0	0.01	0.0	0.01	0.02	0.0	0.95
OtP	0.0	0.0	0.05	0.0	0.02	0.04	0.05	0.11	0.07	0.09	0.58
FinD	0.01	0.01	0.21	0.02	0.06	0.14	0.06	0.28	0.16	0.04	0.0
Out/In	1.66	2.18	1.51	1.94	1.95	3.11	2.16	2.73	2.65	2.15	/
EigC	0.01	0.01	0.18	0.01	0.04	0.08	0.04	0.16	0.10	0.03	0.35
SyncCL	0.01	0.01	0.18	0.01	0.04	0.07	0.04	0.17	0.10	0.03	0.36
SyncCH	0.01	0.00	0.17	0.01	0.04	0.07	0.04	0.16	0.09	0.02	0.39

Note: AtB: Agriculture; C: Mining; D: Manufacturing; E: Utilities; F: Construction; GtH: Trade; I: Transports; JtK: Finance and other business services; LtN: Government services; OtP: Other services; FinD: Final demand.

This table tells a similar story to what could be understood visually in Figure 4.1. First, some sectors produce output that is mostly directly consumed (or invested), rather than used as an intermediate good for further production processes (LtN-Government sends 95%

of its output to final demand). Other sectors, instead, sell most of their output to other production sectors. Second, out of the production sectors, manufacturing is by far the sector which receives most output from other sectors.

From this table, it is not obvious which sectors have a strongest influence on determining the common frequency of oscillation, in case phase synchronization is obtained. One could argue that the size of sectors (i.e., their share of output) should be directly proportional to this importance. For example, it could be that "JtK - Finance and other business services" is more important than "D - Manufacturing" in determining the frequency of oscillations, because the business sector is larger than the manufacturing sector.

However, sectors are heterogeneous in terms of the output/input ratio. This ratio is higher for sectors with higher value added, i.e. sectors that rely less on intermediate goods as an input, relying instead on labor and capital (and potentially producing more profits). Unsurprisingly, "D - Manufacturing" has the lowest value of the output/input ratio, producing about \$1.5 dollars of output for each \$1 dollar of input of intermediate goods. The reason is that at each step of the value chain, value added is small as compared to the value of the intermediate goods. This is different is service sectors. "GtH - Trade" has the highest value of the output/input ratio, 3.11. Since industrial sectors are more connected in the input-output network, should they have larger importance?

As a counterargument, service sectors are heavily influenced by final demand, and, given their size, have an important influence on it. Since final demand is a central node in the economy, closer connection to final demand could actually mean that service sectors are more important in determining the common frequency of oscillations than what one could predict based on their size and general connectedness.

As a first pass to take these feedback effects into account, we calculate the Eigenvector Centrality of the weighted network represented in Table 5.1, with $N + 1 = 11$ nodes. The input-output table is a row-stochastic matrix with all positive entries, so the first left eigenvector of the matrix gives the Eigenvector Centrality. This also corresponds to the steady state of a Markov Chain defined on this matrix. The results are reported in the line "EigC" in Table 5.1. The most interesting result is that "D - Manufacturing" is more important than "JtK - Business" in terms of this centrality measure, precisely because the higher requirement of intermediate inputs in manufacturing makes many sectors of the economy depend on manufacturing demand.

We now perform simulations to explicitly see which sectors are most important in determining the frequency of oscillations. As in the previous section, we obtain heterogeneous frequencies by assigning different values of the accumulation dislike parameter α_1^i, $\alpha_1^i \in [-0.1, -0.02]$, to different nodes. We implement the following algorithm.

1. We focus on a sector, say "AtB - Agriculture", and we assign it the lowest (L) frequency of oscillation (i.e. $\alpha_1^i = -0.02$).

2. For 1000 times, we randomly assign the other frequencies of oscillation to the remaining sectors.

3. For each of these 1000 times, we simulate the model in Eqs. (4.6), calibrated on the US input-output network, and calculate the common frequency of oscillation (under the frequency spread that we consider and the interaction network based on the US input-output network, phase synchronization is always obtained).

4. We repeat the three steps above focusing on each of the 11 nodes in turn.

5. For each focus sector, we average the common oscillation frequencies across random assignments. We end up with a vector of 11 averaged frequencies obtained by assigning the lowest frequency to any focus sector and averaging across many possible assignments of the other frequencies to the other sectors.

6. We compare these 11 frequencies to a benchmark frequency obtained by simulating the model on a uniform weight matrix, that is a matrix in which all entries are $1/(N+1)$. If any frequency is smaller than the benchmark frequency, it means that the focus sector is more important in setting the common frequency of oscillation than in the uniform benchmark case. If the frequency is larger, it means that it is less important.

7. By taking the difference between the benchmark frequency and each of the 11 frequencies, we obtain positive and negative numbers. We add the absolute value of the minimum number to make all frequencies non-negative, and we normalize them so that they sum to unity. The resulting *synchronization centrality* measure, that we denote *SyncCL* (*L* stands for "Lowest frequency assigned to the focus sector"), is reported in the second-to-last row in Table 5.1.

8. We repeat the above procedure assigning the highest (H) frequency to the focus sector. The corresponding synchronization centrality measure, that we denote *SyncCH*, is reported in the last row in Table 5.1.

As is possible to see, both synchronization centrality measures are strongly correlated to eigenvector centrality. Therefore these simulations support the fact that eigenvector centrality could be taken as a measure of the importance of each node in determining the common frequency of oscillation. This confirms the finding that manufacturing has more power than business services in setting the common frequency, because it is more central in the intermediate good flows of the economy.

A final important point is about the robustness of these results to disaggregation. For example, if, instead of considering an aggregate manufacturing sector, we had considered 11 manufacturing sub-sectors, would we get the same results by ex-post aggregating the sub-sectors? In other words, when is it legitimate to aggregate sub-sectors into macro-sectors and when is it not? Ultimately, when is it legitimate to aggregate individual firms into sectors?

One possible answer was already given in Section 5.3.1. If firms in a given sector, or if sub-sectors that are part of a macro-sector, are hit by common shocks, this synchronizes their frequencies. So one could consider the subsector or the macro-sector directly as the relevant unit of synchronization. However, such explanation is only partly satisfactorily, especially given that, as shown in Figure 5.6, synchronization through common noise is not really strong.

To address this issue, we compare the eigenvector centrality for the 10 macro-sectors (plus final demand) in Table 5.1 to the eigenvector centrality calculated from an equivalent weight matrix with 27 sectors (and final demand). These sectors comprise manufacturing and service subsectors and the smallest sectors in Table 5.1, such as "AtB - Agriculture", "C - Mining" and "E - Utilities". The eigenvector centrality is a vector with 28 elements. Aggregating these back into the 11 elements in Table 5.1, we get that the vector of centrality measures is identical to the one in Table 5.1 up to two significant digits.

This example suggests that the importance of nodes in determining the common frequency of oscillation is independent of the level of aggregation, to the extent that synchronization centrality always corresponds to eigenvector centrality.

Synchronization of international business cycles. We now aggregate all sectors within a country, including final demand, and study the international synchronization of endogenous business cycles. We only focus on 17 advanced economies which had a fairly similar development in the last 60 years. These are the US, Canada, Australia, Japan, and 13 European countries (UK, France, Germany, Italy, Portugal, Spain, Belgium, Austria, Ireland, Netherlands, Denmark, Sweden, Finland).

To start, we consider the equivalent version of Table 5.1 in which, instead of having sectors and final demand, we only have countries. In Table 5.2, each row-column pair indicates the flow of goods and services from row to column. For example, 0.03 at the "AUS" row and "JPN" column indicates that 3% of the Australian output is sold to Japan. The diagonal of this matrix is the fraction of the value flows that stays within a country, in the same way as the diagonal of Table 5.1 was indicating the fraction of value flows that was staying within a sector.

Table 5.2: International Input-Output (Year 2000)

Country	AUS	AUT	BEL	CAN	DEU	DNK	ESP	FIN	FRA	GBR	IRL	ITA	JPN	NLD	PRT	SWE	USA
AUS	0.93	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.0	0.0	0.03	0.0	0.0	0.0	0.02
AUT	0.0	0.64	0.01	0.0	0.17	0.0	0.01	0.0	0.02	0.03	0.0	0.05	0.01	0.01	0.0	0.01	0.03
BEL	0.0	0.01	0.57	0.0	0.1	0.01	0.02	0.0	0.09	0.05	0.0	0.04	0.01	0.05	0.0	0.01	0.04
CAN	0.0	0.0	0.0	0.66	0.01	0.0	0.0	0.0	0.0	0.01	0.0	0.0	0.01	0.0	0.0	0.0	0.29
DEU	0.0	0.01	0.02	0.0	0.8	0.0	0.01	0.0	0.04	0.02	0.0	0.02	0.01	0.01	0.0	0.01	0.03
DNK	0.0	0.0	0.01	0.0	0.06	0.7	0.01	0.01	0.02	0.03	0.01	0.01	0.01	0.01	0.0	0.06	0.04
ESP	0.0	0.0	0.01	0.0	0.03	0.0	0.85	0.0	0.04	0.01	0.0	0.02	0.0	0.01	0.02	0.0	0.01
FIN	0.0	0.01	0.01	0.01	0.07	0.01	0.02	0.67	0.03	0.05	0.0	0.02	0.01	0.02	0.0	0.04	0.03
FRA	0.0	0.0	0.02	0.0	0.04	0.0	0.02	0.0	0.83	0.02	0.0	0.02	0.0	0.01	0.0	0.0	0.02
GBR	0.0	0.0	0.01	0.01	0.04	0.0	0.01	0.0	0.02	0.82	0.02	0.01	0.01	0.01	0.0	0.01	0.04
IRL	0.0	0.0	0.02	0.01	0.08	0.0	0.01	0.0	0.04	0.09	0.49	0.02	0.02	0.01	0.0	0.01	0.18
ITA	0.0	0.0	0.01	0.0	0.03	0.0	0.01	0.0	0.03	0.01	0.0	0.87	0.0	0.0	0.0	0.0	0.02
JPN	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.96	0.0	0.0	0.0	0.02
NLD	0.0	0.01	0.06	0.0	0.09	0.01	0.01	0.0	0.04	0.05	0.01	0.03	0.01	0.62	0.0	0.01	0.04
PRT	0.0	0.0	0.01	0.0	0.03	0.0	0.04	0.0	0.02	0.02	0.0	0.01	0.0	0.01	0.83	0.0	0.01
SWE	0.0	0.01	0.02	0.01	0.06	0.04	0.01	0.02	0.02	0.04	0.0	0.02	0.01	0.02	0.0	0.7	0.04
USA	0.0	0.0	0.0	0.02	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.0	0.0	0.0	0.95
Output	0.02	0.01	0.01	0.03	0.08	0.01	0.03	0.01	0.06	0.06	0.00	0.05	0.20	0.02	0.00	0.01	0.42
Out/In	1.96	2.21	1.83	2.12	2.13	2.22	2.01	1.96	2.09	2.10	1.82	1.95	2.15	1.99	1.99	2.01	2.20
EigC	0.01	0.01	0.01	0.03	0.08	0.01	0.04	0.00	0.07	0.06	0.00	0.06	0.13	0.01	0.01	0.01	0.47

The difference is immediate. While most of what was produced within a sector was flowing to other sectors, most of what is produced within a country stays within the same country. This ranges from 96% in the case of Japan and 95% in the case of the US, to 49% in the case of Ireland. On average, 90% of the value flows stay within the same country.[69]

In Table 5.2, we also report the shares of output, the output/input ratio and the Eigenvector Centrality. Interestingly, the output/input ratio does not vary much, differently from the case of sectors, indicating that the share of value added is similar across countries.[70] The Eigenvector Centrality is strongly correlated to the shares of output. The US have Eigenvector Centrality 0.47, while the European countries, combined, have 0.35. This shows that the US have most power in setting the rhythms of global business cycles, and quantifies this power.

We now simulate our model on the international input-output network. In Figure 5.7, we consider the usual frequency spread, induced by heterogeneous $\alpha_1^i \in [-0.1, -0.02]$, which was leading to synchronization in the US input-output network. As is clear from the figure, in the international input-output network phase synchronization is no longer obtained. Different lines indicate different countries, and colors are scaled in a way that blue corresponds to high frequencies and red corresponds to low frequencies. Frequencies are assigned

[69] While these data refer to year 2000, when globalization had not completely taken off yet, Cerina et al. (2015) show that the foreign share of value flows had only increased to about 12.5% in 2011, after a small dip corresponding to the 2008 financial crisis. Likewise, the foreign share in year 1995 was only 1% smaller than in year 2000.

[70] This ratio would probably be smaller in countries like China, which base their economy on manufacturing much more, but within the advanced economies in our sample, the input-output ratio is roughly constant.

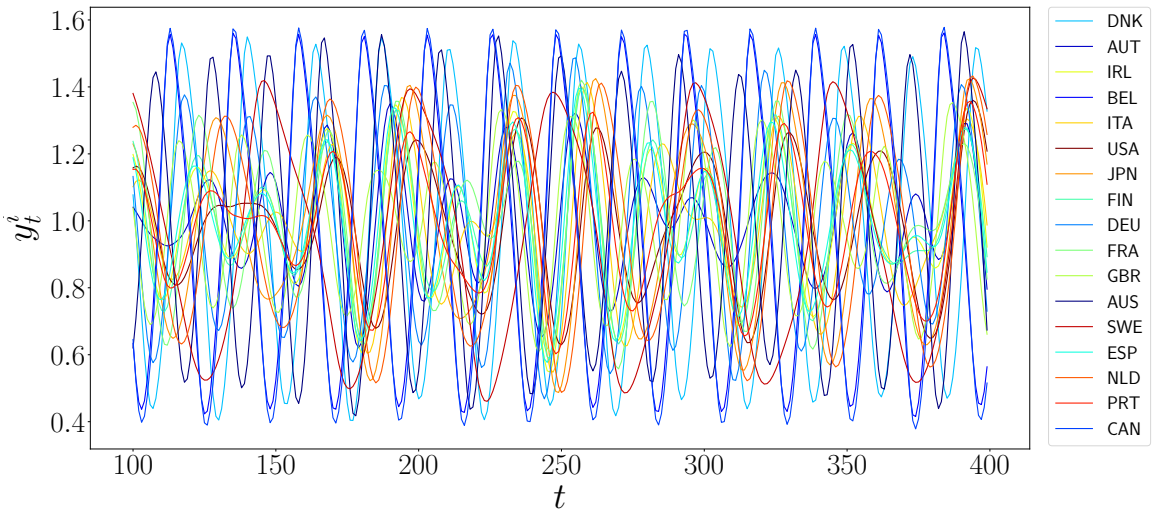
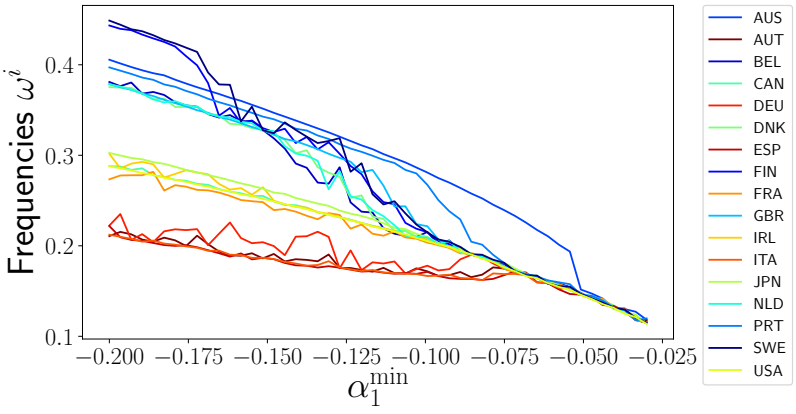


Figure 5.7: Time series generated by model (4.6), calibrated on the international input-output network. Frequencies of oscillation are randomly assigned to different countries. As in the rest of this chapter, blue (red) lines are associated to higher (lower) frequencies.

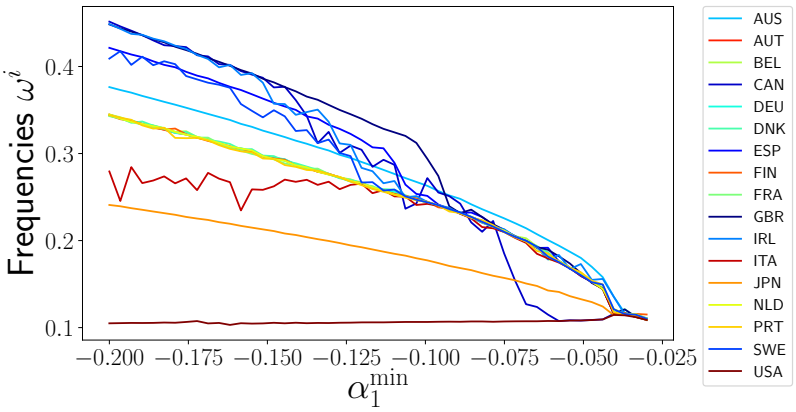
randomly to countries, as it is not clear which countries should experience faster vs slower oscillations. Looking at the figure, some countries that have middle to low frequencies (from green to red) have formed a cluster of almost synchronized nodes, while another cluster at high frequencies (in light and dark blue) is also present. The US, which, according to the last row of Table 5.2, have by far the highest Eigenvector Centrality, are part of the first cluster, and are likely the key driver that holds it together. The other cluster is formed by a number of European countries and Australia, which are apparently not as dependent on the US and have similar frequencies.

We now study a scenario of path to synchronization (Gómez-Gardenes et al., 2007; Arenas et al., 2008; Navas et al., 2015) as the frequency spread is reduced from that induced by a minimum value of α_1^i, $\alpha_1^{\min} = -0.2$, so that the range of parameters α_1^i vary in $\alpha_1^i \in [-0.2, -0.02]$, to that corresponding to a minimum $\alpha_1^{\min} = -0.03$, so that the parameters $\alpha_1^i \in [-0.03, -0.02]$. Clearly, synchronization is harder if frequency spread is larger. One can expect synchronization to be possible for the minimum spread, but exactly how synchronization is obtained going from the largest to the smallest frequency spread depends on the network structure, highlighting key structural properties of the interactions.

In Figure 5.8, we study the path to synchronization for two non-representative random assignments of frequencies to different countries. In Figure 5.8a, the US are assigned an intermediate frequency. The US form a synchronized cluster of intermediate frequencies which, as the frequency spread is reduced, incorporates the other clusters at extreme values of the frequencies. Australia, although its frequency is not as high as that of Sweden or Belgium (the blue line corresponding to Australia is less dark), is the last country to adapt



(a)



(b)

Figure 5.8: Path to synchronization as the frequency spread is reduced from that induced by $\alpha_1^{\min} = -0.2$ to that induced by $\alpha_1^{\min} = -0.03$, for two non-representative random assignments of frequencies to different countries.

its frequency to the other countries, due to its trade isolation. The situation in Figure 5.8b is similar in terms of the formation of synchronized clusters that merge as the frequency spread is reduced, until global synchronization is obtained for minimum frequency spread. The US, which in this case are assigned the lowest frequency, are the last country that joins the main cluster. Note the blue line that converges to the US line at $\alpha_1^{\min} = -0.06$. This is Canada, which, being very closely connected to the US, and being much smaller than it, adjusts its rhythm to the US frequency before adjusting to the main cluster.

These two examples were valid for two specific random assignments of frequencies. We now test the properties of the model as the frequency spread is increased in a way that is independent of the specific random assignment of frequencies. For each value of $\alpha_1^{\min}$, we simulate the model for 100 different random assignments, and then we average the results. In Figure 5.9 we show how three quantities vary with $\alpha_1^{\min}$. These are phase coherence and the average correlation coefficient, already used before, and the frequency spread. This latter

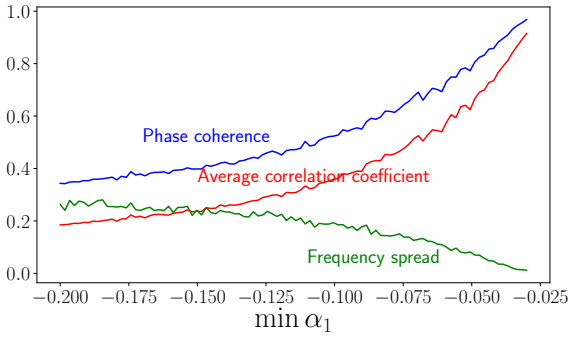


Figure 5.9: Average correlation coefficient, phase coherence and frequency spread for $\alpha_1^{\min} \in [-0.2, -0.03]$, averaging over 100 random assignments of frequencies to different countries.

quantity is defined as the ratio between the standard deviation and the mean of frequencies (it is zero in case of frequency entrainment, because all frequencies are the same). As the figure shows, decreasing $\alpha_1^{\min}$ makes the dynamics mostly incoherent (uncorrelated) and increases the frequency spread.

Finally, we discuss the interplay between sectors and countries. Some sectors are quite internationally open, while other sectors are not. In the following, we give the national share of value flows for each sector, averaged across countries: AtB - Agriculture, 88%; C - Mining, 77%; D - Manufacturing, 73%; E - Utilities, 99%; F - Construction, 100%; GtH - Trade, 96%; I - Transport, 89%; JtK - Business, 96%; LtN - Government, 100%; OtP - Other Services, 98%. This shows that manufacturing, mining and agriculture are the most internationally open, while service sectors have very small openness. Therefore, given the small size of agriculture and mining, international synchronization of endogenous business cycles is likely to occur through manufacturing. In a scenario of path to synchronization, first manufacturing sectors in different countries would synchronize, and then this would get transmitted to the rest of the economy. We will explore the role of manufacturing as a "bridge" across countries thoroughly in terms of complete synchronization in the next section.

5.4 Complete synchronization

All analysis of phase synchronization so far was concerned with the conditions under which the frequencies of oscillation of different agents could become similar. This analysis is limited by the extent to which we can empirically measure the uncoupled frequency of a sector or a country. It is also limited by the assumption that endogenous business cycles can be described by non-linear dynamic models that exhibit a relatively clear periodicity. One may want to consider a high-dimensionally chaotic business cycle model, and phase

synchronization would be of little use. In this section, we study synchronization from a different perspective. We assume that the frequencies of oscillation are identical, and we explore the conditions under which complete synchronization can emerge, i.e. the non-linear dynamics of the agents become perfectly aligned. Since we are studying a limit cycle model, complete synchronization is always obtained absent shocks. Yet, the theory of complete synchronization that we develop can be used to study the interplay between synchronizing non-linear dynamics, exogenous shocks and network structure.

This is particularly useful to explain how synchronization theory could improve on existing exogenous business cycle models in terms of generating higher comovement between economic time series. Consider two agents following two deterministic identical limit cycles in isolation. When coupled, the results that we will present show that their dynamics perfectly align. Denote this as a *synchronized state* $\boldsymbol{s}_t = (x_t^s, y_t^s)$, such that $x_t^i = x_t^s$ and $y_t^i = y_t^s$, for $i = 1, 2$. Suppose now that at time t, the decision variables y^i of the agents are hit by idiosyncratic shocks ξ_t^1 and ξ_t^2. The comovement between the time series of the decision variables, calculated for the time steps immediately following the shocks, is given by the Pearson correlation coefficient

$$\operatorname{cor}\left(y_t^s + \xi_t^1, y_t^s + \xi_t^2\right) = \frac{\operatorname{cov}\left(y_t^s, y_t^s\right) + \operatorname{cov}\left(y_t^s, \xi_t^2\right) + \operatorname{cov}\left(\xi_t^1, y_t^s\right) + \operatorname{cov}\left(\xi_t^1, \xi_t^2\right)}{\operatorname{std}\left(y_t^s + \xi_t^1\right) \cdot \operatorname{std}\left(y_t^s + \xi_t^2\right)}. \tag{5.7}$$

In the above equation, cov denotes the covariance, std indicates the standard deviation, and we have used linearity of the covariance to decompose it in various terms. Compare this equation with the correlation between time series of two agents that are in stable steady state that is hit by shock (ξ_t^1, ξ_t^2)

$$\operatorname{cor}\left(\xi_t^1, \xi_t^2\right) = \frac{\operatorname{cov}\left(\xi_t^1, \xi_t^2\right)}{\operatorname{std}\left(\xi_t^1\right) \cdot \operatorname{std}\left(\xi_t^2\right)}. \tag{5.8}$$

If we assume that the variance of endogenous and exogenous fluctuations is the same (i.e. fluctuations have the same intensity), the correlation coefficient is only determined by the numerators of these expressions. Comparing Eq. (5.7) and Eq. (5.8), it is clear that if the shocks hitting the synchronized state y_t^s are small relative to the limit cycle, correlation is larger in the endogenous case. (This is confirmed by numerical simulations.) In fact, if shocks are really small, the correlation coefficient approaches one.

In the following, we explore how shocks propagate on the interaction network on the path back to the (stable) synchronized state. As we did in Section 5.3, we first study complete synchronization in an abstract interaction network (Section 5.4.1), and then we consider input-output networks (Section 5.4.2).

5.4.1 Abstract interaction network

To study the interplay between synchronizing dynamics, shock propagation and network structure, we adapt the master stability approach originally proposed by Pecora and Carroll (1998). We start by assuming that all agents are in the synchronized state $\boldsymbol{s}_t = (x_t^s, y_t^s)$, such that $x_t^i = x_t^s$ and $y_t^i = y_t^s$, $\forall i$. This could have been reached after an initial transient. The starting point of the master stability approach is to calculate the relative strength of desynchronizing forces that lead to deviations from the synchronized state, and synchronizing forces due to coupling. Here, desynchronizing forces only correspond to idiosyncratic shocks u_t^i, because limit cycles are attractive. However, in the case of chaotic dynamical systems, chaos acts as a desynchronizing force, too. The key idea of the master stability approach is to perform a transformation that uncouples the deviations from the synchronized state in orthogonal components. The approach is semi-analytical, in the sense that the relative strength of synchronizing and desynchronizing forces is obtained by numerically computing Lyapunov exponents, but once this is computed, the key insights are obtained analytically.

Let $x_t^i = x_t^s + x_t^{\xi,i}$ and $y_t^i = y_t^s + y_t^{\xi,i}$. Here, $x_t^{\xi,i}$ and $y_t^{\xi,i}$ denote a small deviation from the synchronized state, for example due to idiosyncratic noise hitting sector i. Let further $\boldsymbol{\xi}_t = \left(x_t^{\xi,1}, y_t^{\xi,1}, x_t^{\xi,2}, y_t^{\xi,2}, \dots x_t^{\xi,N}, y_t^{\xi,N}\right)$ be the $2N$-dimensional vector of deviations.

To show how the evolution of this vector can be obtained, we fully work out a specific example. Consider a star network with three nodes. This is composed by a central node, which we denote as node 1, that is linked to two leaf nodes, which we denote as nodes 2 and 3. The leaf nodes are only connected to the central node. Replacing $x_t^i = x_t^s + x_t^{\xi,i}$ and $y_t^i = y_t^s + y_t^{\xi,i}$ in Eq. (4.2) and assuming, as we did in Section 5.3.1, that the dependence on other nodes is ϵ/K_i, where K_i is the degree of node i, while the dependency on oneself is $1-\epsilon$, we get the following system of equations for the evolution of $\boldsymbol{\xi}_t$:

$$\begin{cases} x_{t+1}^{\xi,1} = (1-\delta)x_t^{\xi,1} + y_t^{\xi,1}, \\ y_{t+1}^{\xi,1} = \alpha_1 x_t^{\xi,1} + (\alpha_2 + F'(y_t^s))\, y_t^{\xi,1} - \epsilon F'(y_t^s)\left(y_t^{\xi,1} - \frac{1}{2}y_t^{\xi,2} - \frac{1}{2}y_t^{\xi,3}\right), \\ x_{t+1}^{\xi,2} = (1-\delta)x_t^{\xi,2} + y_t^{\xi,2}, \\ y_{t+1}^{\xi,2} = \alpha_1 x_t^{\xi,2} + (\alpha_2 + F'(y_t^s))\, y_t^{\xi,2} - \epsilon F'(y_t^s)\left(y_t^{\xi,2} - y_t^{\xi,1}\right), \\ x_{t+1}^{\xi,3} = (1-\delta)x_t^{\xi,1} + y_t^{\xi,1}, \\ y_{t+1}^{\xi,3} = \alpha_1 x_t^{\xi,1} + (\alpha_2 + F'(y_t^s))\, y_t^{\xi,1} - \epsilon F'(y_t^s)\left(y_t^{\xi,3} - y_t^{\xi,1}\right), \end{cases} \tag{5.9}$$

where we have used the fact that the deviations $\boldsymbol{\xi}_t$ are small, to Taylor-expand the function F. This formulation suggests that the evolution of the deviations of any single node is given by the Jacobian of the dynamics corresponding to the node, plus a term of interaction with

other nodes. Extrapolating from this example, it is easy to see that the evolution of the whole vector $\boldsymbol{\xi}_t$ is generally given by

$$\boldsymbol{\xi}_{t+1} = \left(\boldsymbol{I}_N \otimes \boldsymbol{J}(\boldsymbol{s}_t) - \epsilon F'(\boldsymbol{s}_t)\boldsymbol{KLK} \otimes \boldsymbol{H}\right)\boldsymbol{\xi}_t, \tag{5.10}$$

where

- $\boldsymbol{I}_N$ is the N-dimensional identity matrix.
- $\boldsymbol{J}(\boldsymbol{s}_t)$ is the 2-dimensional Jacobian.
- $\otimes$ denotes the Kronecker product, i.e. a $2N$-dimensional matrix.
- $F'(\boldsymbol{s}_t)$ is the first derivative of F, due to the fact that we are considering a first-order approximation around the synchronized state.
- $\boldsymbol{K}$ is an N-dimensional square matrix with $1/\sqrt{k_i}$ on the main diagonal and zero everywhere else.
- $\boldsymbol{L}$ is the Laplacian of the network. This is a key mathematical property of the network that is widely used in many applications. $L_{ii} = k_i$ and $L_{ij} = -1$ if i and j are connected. $\boldsymbol{KLK}$ is known as normalized Laplacian, and has similar properties to the Laplacian.[71]
- $\boldsymbol{H}$ is the 2-dimensional square matrix of connectivity in the dynamical system. Since here the only connectivity is through y, it is $H_{22} = 1$ and zero everywhere else.

The problem with Eq. (5.10) is that the evolution of each component $y_t^{\xi,i}$ depends on all other agents j to which i is connected. The key trick of the master stability approach is to diagonalize $\boldsymbol{KLK}$ so to decompose the deviations in orthogonal, uncoupled, components. Following the terminology in the literature, we call these components *eigenmodes*. Let $\boldsymbol{\xi}_t = (\boldsymbol{Q} \otimes \boldsymbol{I}_2)\,\boldsymbol{\zeta}_t$, where Q is the matrix of eigenvectors of $\boldsymbol{KLK}$. Here, $\boldsymbol{\zeta}_t$ can be interpreted as a projection of $\boldsymbol{\xi}_t$ in the eigenspace. It evolves according to

$$\boldsymbol{\zeta}_{t+1} = \left(\boldsymbol{I}_N \otimes \boldsymbol{J}(\boldsymbol{s}_t) - \epsilon F'(\boldsymbol{s}_t)\boldsymbol{\Lambda} \otimes \boldsymbol{H}\right)\boldsymbol{\zeta}_t, \tag{5.11}$$

[71] Here we summarize a few properties of $\boldsymbol{KLK}$. Because the rows sum to zero, one eigenvalue is zero. Moreover, it is well known that the other eigenvalues are positive and bounded between 0 and 2. The multiplicity of the 0 eigenvalue reflects the number of disconnected clusters in the network. We sort the eigenvalues in increasing order, so that $\lambda_1 = 0$, $\lambda_2 > 0$ if the network is connected, $\lambda_3 \geq \lambda_2$, ..., $\lambda_N \geq \lambda_{N-1}$, $\lambda_N \leq 2$. The eigenvalue λ_2 is knows as *algebraic connectivity* and the corresponding eigenvector as *Fiedler vector*. The smaller λ_2, the more the network has a modular structure, in which two or more clusters of nodes have strong internal connectivity and weak external connectivity. In the case of two clusters, the components of the Fiedler vector are positive for nodes in a cluster and negative for nodes in the other cluster. Fiedler vectors are commonly used for graph partitioning.

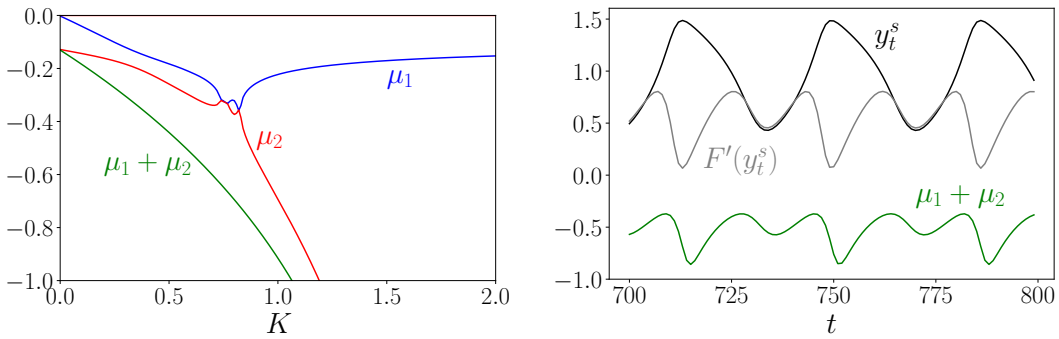
where $\mathbf{\Lambda}$ is the matrix with the eigenvalues of $\boldsymbol{KLK}$ on the main diagonal and zero everywhere else. Each of the N eigenmodes of $\boldsymbol{\zeta}_t$ can then be written as

$$\boldsymbol{\zeta}^i_{t+1} = \left(\boldsymbol{J}(\boldsymbol{s}_t) - F'(\boldsymbol{s}_t)\epsilon\lambda_i \boldsymbol{H}\right)\boldsymbol{\zeta}^i_t. \tag{5.12}$$

In this basis, the evolution of each eigenmode i only depends upon itself. We stress that i now corresponds to an eigenvector (of the eigenvalue λ_i of the normalized Laplacian of the network) and not to an agent. As will be clear in the examples below, the evolution of each eigenmode corresponds to higher-order properties of the network. Of course one can retrieve the dynamics of the agents by applying the transformation $\boldsymbol{\xi}_t = (\boldsymbol{Q} \otimes \boldsymbol{I}_2)\,\boldsymbol{\zeta}_t$.

The eigenmode $i = 1$ corresponds to dynamics parallel to the synchronization manifold (i.e., "in the same direction as the dynamics"). It captures the phase shift due to a shock hitting the system at a given time t. The eigenmodes $i > 1$ correspond to dynamics orthogonal to the synchronization manifold. If these dynamics always converge to zero after the shock, the synchronized state is stable.

Whether the orthogonal dynamics converge to zero is not obvious from Eq. (5.12). To know if they do, one must numerically compute the Lyapunov exponents of the system. Letting $K = \epsilon\lambda_i$ denote an effective coupling for eigenmode i, one calculates Lyapunov exponents in Eq. (5.12) for all values of K that are typically obtained. In our case, because $\epsilon \in [0, 1]$ and $\lambda_i \in [0, 2]$, it is of interest to only consider the interval $K \in [0, 2]$.



(a) (b)

Figure 5.10: (Left) Time-averaged Lyapunov exponents μ_1 and μ_2, and sum $\mu_1 + \mu_2$, as a function of the effective coupling K. (Right) Time evolution of the sum of Lyapunov exponents $\mu_1 + \mu_2$, compared to the time evolution of y at the synchronized state y^s_t, and to the time evolution of $F'(y^s_t)$.

The Lyapunov exponents are shown in Figure 5.10, for a parameterization of the model that we will specify below. In Figure 5.10a we only consider the time average of the Lyapunov exponents. The dependence of the first Lyapunov exponent μ_1 on the effective coupling K is called *master stability function*. The first thing that stands out is that both μ_1 and μ_2 are always negative. This means that perturbations along the orthogonal eigenmodes indeed vanish over time. If the system was chaotic, this would no longer be true, and

the synchronized state would only be stable for certain values of K. Moreover, the largest Lyapunov exponent μ_1 is null for $K = 0$: this means that perturbations in the direction parallel to the synchronization manifold do not vanish, and there is indeed a permanent phase shift. Finally, the highest values of μ_1 and μ_2 are obtained for smallest K. All else equal, it means that eigenmodes corresponding to the smallest eigenvalues λ_i are the ones that decay most slowly, while eigenmodes corresponding to relatively large eigenvalues decay much more rapidly.

As in all dynamical systems, the Lyapunov exponents are never constant, but rather vary along the attractor. Figure 5.10b shows the time evolution of the sum of Lyapunov exponents $\mu_1 + \mu_2$, representing the rate of expansion or contraction of volumes in phase space. The times at which $\mu_1 + \mu_2$ is minimal (most negative) are those at which the effects of shocks are weakest and most rapidly absorbed, while maximal (least negative) $\mu_1 + \mu_2$ indicates that shocks have a stronger and more lasting effect. To understand to which phases of the evolution of the system these correspond, we plot the dynamics of the synchronized state y_t^s and the time evolution of $F'(y_t^s)$. It is interesting to note that the evolution of $\mu_1 + \mu_2$ closely mirrors that of $F'(y_t^s)$. The most stable situation is when y_t^s is in a peak, followed by y_t^s being in a trough. The most unstable situations correspond instead to periods of rapid growth or recession. In this situations, shocks propagate most strongly and lead to strongest desynchronization.

In summary, our theory calculates the deviations from a synchronized state in which all agents behave alike—deviations that could for example be due to an exogenous shock hitting one agent only. This is an approximation due to the fact that we assume that deviations are small to consider a Taylor expansion of the function F. Instead of calculating the evolution of the deviations for each agent, we calculate the evolution of deviations for each eigenmode, corresponding to an eigenvector of a matrix of connectivity. This allows (i) to show that the synchronized state is stable to small shocks, but it is generally affected by a permanent phase shift; (ii) to understand how the propagation of the shock is affected by the interaction network; (iii) to understand how the propagation of the shock is affected by the timing of the shock with respect to the endogenous dynamics. Ultimately, this theory is useful to assess which level of correlation we can expect between economic time series that are modeled by our system. To get more intuition into this formalism, we now consider a number of examples.

All parameters in the examples below are chosen following the guidelines in Section 5.2. The values of β_0, β_1, β_2, β_3, β_4, δ and α_0 are chosen as in Section 5.2. We further choose $\alpha_2 = 0.4$ and $\alpha_1 = -0.04$, which leads to a cycle of about 36 time units. If one time unit

corresponds to a quarter, this cycle would have length 9 years, which is in line with the empirical evidence of employment variables in the United States (Beaudry et al., 2019).

In this section, we consider a single idiosyncratic shock hitting the agents of the network at time τ, and study its propagation. In the next section, while considering the empirical input-output network, we analyze the effect of idiosyncratic shocks repeatedly hitting sectors and countries.

Two connected agents. We start from the simplest possible system to study the effect of the dynamics and the shocks in most detail. We will then consider a more interesting interaction structure to understand shock propagation. Here, we consider two connected agents, giving weight $1-\epsilon$ to their own dynamics and ϵ to the dynamics of the other node. We assume $\epsilon = 0.3$. The eigenvalues of $\boldsymbol{KLK}$ are $\lambda_1 = 0$ and $\lambda_2 = 2$, the corresponding eigenvectors are $\boldsymbol{v}_1 = (1,1)/\sqrt{2}$ and $\boldsymbol{v}_2 = (-1,1)/\sqrt{2}$. After the shock hits at time τ, at time $t \geq \tau$ the deviations from the synchronized state in the initial basis and in the eigenbasis are related by the following equations:

$$\begin{aligned}
x_t^{\xi,1} &= \tfrac{1}{\sqrt{2}}\left(x_t^{\zeta,1} - x_t^{\zeta,2}\right), & x_t^{\zeta,1} &= \tfrac{1}{\sqrt{2}}\left(x_t^{\xi,1} + x_t^{\xi,2}\right), \\
y_t^{\xi,1} &= \tfrac{1}{\sqrt{2}}\left(y_t^{\zeta,1} - y_t^{\zeta,2}\right), & y_t^{\zeta,1} &= \tfrac{1}{\sqrt{2}}\left(y_t^{\xi,1} + y_t^{\xi,2}\right), \\
x_t^{\xi,2} &= \tfrac{1}{\sqrt{2}}\left(x_t^{\zeta,1} + x_t^{\zeta,2}\right), & x_t^{\zeta,2} &= \tfrac{1}{\sqrt{2}}\left(-x_t^{\xi,1} + x_t^{\xi,2}\right), \\
y_t^{\xi,2} &= \tfrac{1}{\sqrt{2}}\left(y_t^{\zeta,1} + y_t^{\zeta,2}\right), & y_t^{\zeta,2} &= \tfrac{1}{\sqrt{2}}\left(-y_t^{\xi,1} + y_t^{\xi,2}\right).
\end{aligned} \tag{5.13}$$

The first column of Eq. (5.13) already highlights an interesting pattern. Given our previous discussion, the second eigenmode $\left(x_t^{\zeta,2}, y_t^{\zeta,2}\right)$ eventually converges to zero, because $K = \epsilon\lambda_2 = 0.6 > 0$. The larger the coupling coefficient ϵ, the faster the second eigenmode converges to zero. Independently of speed, eventually both $\left(x_t^{\xi,1}, y_t^{\xi,1}\right)$ and $\left(x_t^{\xi,2}, y_t^{\xi,2}\right)$ converge to the first eigenmode $\left(x_t^{\zeta,1}, y_t^{\zeta,1}\right)$. This shows analytically that, after some time, the deviations of the two agents coincide with the perturbation parallel to the synchronization manifold, reinstating a synchronized state that is different from the initial synchronized state by a constant phase shift. This is true for every interaction structure, as the eigenvector corresponding to $\lambda_1 = 0$ is always of the type $\boldsymbol{v}_1 = (1,\ldots,1)/\sqrt{N}$, and the eigenmodes corresponding to the other eigenvalues eventually vanish.

To proceed, we assume that the single shock at time τ only hits the variables y, i.e. $y_\tau^{\xi,1} = u_1$, $y_\tau^{\xi,2} = u_2$ and $x_\tau^{\xi,1} = x_\tau^{\xi,2} = 0$. From the second column of (5.13) we can see that if $u_1 = u_2$, $y_{\tau+1}^{\zeta,2} = 0$, and so for all following times t, $y_t^{\zeta,2} = 0$: The eigenmode transverse to the synchronization manifold never takes positive values. In other words, because the shock is common, it does not cause any relative difference in phase between the two agents. On the contrary, $y_{\tau+1}^{\zeta,1} = \sqrt{2}u_1$ is maximal: the effect of the shock is to only shift the phase.

If $u_1 > 0$, the phase is shifted forward; if $u_1 < 0$, it is shifted backwards. Consider now the reverse case $u_1 = -u_2$. Now $y^{\zeta,1}_{\tau+1} = 0$ and $y^{\zeta,2}_{\tau+1} = \sqrt{2}u_1$. In this case the transverse eigenmode is maximal, while the parallel eigenmode is null. After a transient in which the phases are different, they synchronize again, with no consequence of the shock in terms of permanent phase shift.

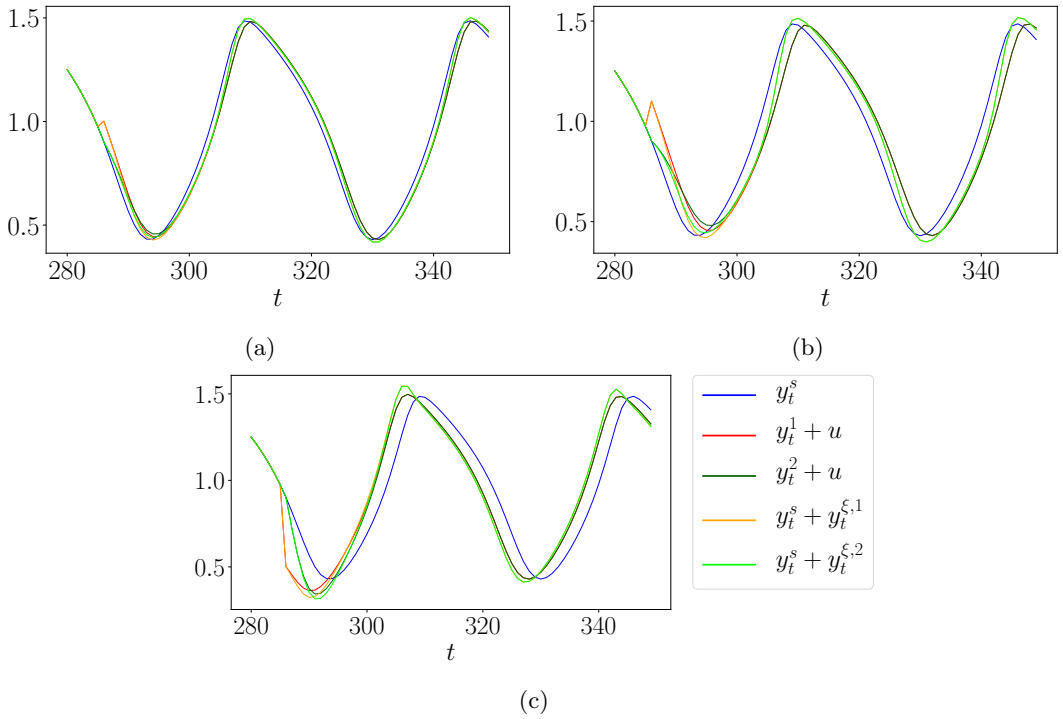


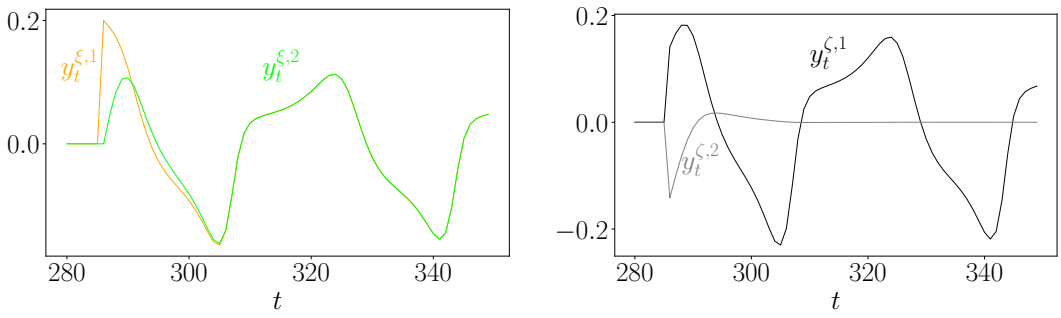
Figure 5.11: Effect of a shock hitting agent 1 only. Blue lines corresponds to y^s_t, the (unperturbed) synchronized state. Red and dark green lines correspond to the actual perturbed dynamics, which we denote as $y^1_t + u$ and $y^2_t + u$ respectively, fully simulated from dynamical system (4.2). The orange and light green lines correspond to the approximation of the perturbed dynamics, as obtained by Eq. (5.10). In panel (a), $u = 0.1$; in panel (b), $u = 0.2$; in panel (c), $u = -0.4$.

We now consider the case in which the shock hits one agent only. Without loss of generality, we set $u_1 = u$ and $u_2 = 0$. In Figure 5.11, we show some simulations of the dynamics, for three values of the shock u: $u = 0.1$ (Fig. 5.11a); $u = 0.2$ (Fig. 5.11b); $u = -0.4$ (Fig. 5.11c).

There are two patterns that are worth noting in Figure 5.11. The first is how good an approximation is $y^s_t + y^{\xi,i}_t$ to $y^i_t + u$. The approximation is excellent for $u = 0.1$ and $u = -0.4$, but it is not as good in the case of $u = 0.2$. Note that these shocks are up to 40% of the steady state value, so the assumption of small shocks that we needed to Taylor expand the function $F(\cdot)$ definitely does not hold. Nonetheless, the approximation is still quite good. Moreover, across all cases the approximation is better at certain times. In particular, it tends to be worse in the final part of the economic boom and in the final part

of the recession. As shown in Figure 5.10b, these correspond to times when $F'(y_t^s)$ changes rapidly, suggesting that at those times one should also consider second order terms $F''(y_t^s)$. The reason why the approximation is better for $u = -0.4$ than for $u = 0.2$, despite the first shock being twice as large in absolute value than the second one, is also due to the change in $F'(y_t^s)$. Indeed, a positive shock hitting during a recession changes the properties of the dynamics much more than a negative shocks does, leading to a worse approximation. A negative shock hitting during a boom has the same effect as a positive shock hitting during a recession.

The other pattern that is worth noting in Figure 5.11 is the fact that a negative shock makes the perturbed dynamics lagging with respect to the unperturbed one, while positive shocks have the opposite effect. This result was already expected given the discussion above, but Figure 5.11 gives a numerical confirmation.

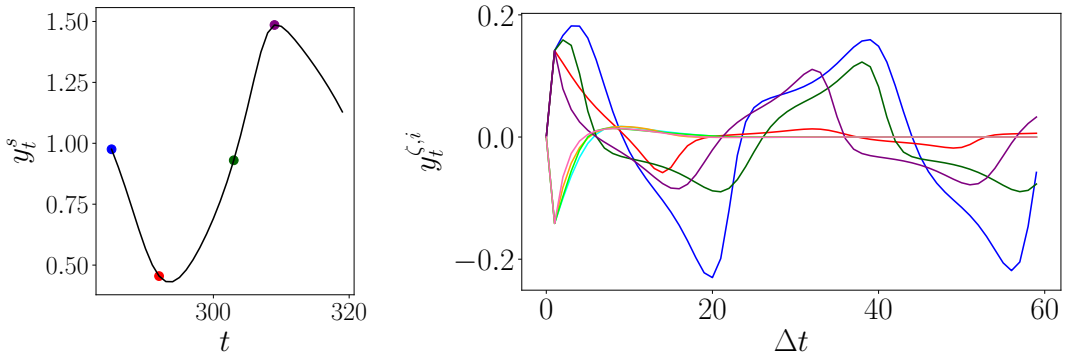


(a) (b)

Figure 5.12: Deviations from the synchronized state when a positive shock $u = 0.2$ hits agent 1 only. Panel (a) shows the deviations in the original basis, while panel (b) shows deviations in the eigenbasis.

In Figure 5.12 we show the evolution of the deviations from the synchronized state in the case of $u = 0.2$. In Figure 5.12a, we show the deviations in the original basis, $y^{\xi,1}$ and $y^{\xi,2}$. As agent 2 did not receive any shock, its dynamics slowly adjusts to that of agent 1, and after a short transient both agents get synchronized again. Figure 5.12b shows the deviations in the eigenbasis. (Note that, because of the normalization of the eigenvectors, the initial shock is in this basis $0.2/\sqrt{2}$.) It is clear that the deviation orthogonal to the synchronization manifold (grey line) quickly vanishes, and only the deviation parallel to the synchronization manifold (black line) persists. This indicates a permanent phase shift.

Finally, in Figure 5.13 we investigate the effect of the timing of shocks. The blue and cyan lines correspond to the same shock $u = 0.2$ at time $t = 285$, for the perturbations parallel and perpendicular to the synchronization manifold respectively. The same is true for the red and orange lines ($t = 292$), the dark green and light green lines ($t = 303$) and purple and pink lines ($t = 309$). As discussed above, the shocks have strongest desynchronizing



(a) (b)

Figure 5.13: Effect of the timing of shocks. Panel (b) shows deviations in the eigenbasis for four times at which the shock could hit: a recession (blue and cyan); a trough (red and orange); a growth period (dark and light green); a peak (purple and pink).

effects during a recession or during a growth period. The desynchronizing effect in a trough is smaller, while the effect at a peak is minimal. It is interesting that most of the differences are due to the first eigenmode, while the evolution of the second eigenmode is almost independent of timing.

It is also interesting that the strongest desynchronizing effect is during a recession rather than during a boom (blue vs dark green line). This can be explained from Figure 5.10b, noting that after a recession is a trough, where the Lyapunov exponent is smaller but not minimal; on the other hand, after a boom is a peak, where the Lyapunov exponent is minimal. Even if the shock occurs at times in which the Lyapunov exponent is identical, the evolution of the deviation is affected differently in the two cases. The dark green line shows how the shock is suddenly dampened, corresponding to the dynamics reaching the peak. The blue line, on the other hand, shows that the shock is not as much dampened, because the dynamics reaches the trough where the Lyapunov exponent is not minimal.

Six agents with two cliques of three nodes each. To understand the effect of the network structure on the propagation of the shock and the synchronizing dynamics, we consider the network in Figure 5.14. This network is composed of two cliques of three nodes each, connected by a link between two of the nodes of the cliques. These two nodes are colored lighter to highlight the stronger connectivity to the other clique. This network can be thought of as a stylized model of sectors in two countries, where sectors within the same country are more connected than across countries. The sectors that connect the two countries can be thought of as the respective manufacturing sectors which, as discussed in Section 5.3.2, act as "bridges" that propagate shocks across countries. We consider this example explicitly in the following input-output interpretation of the model.

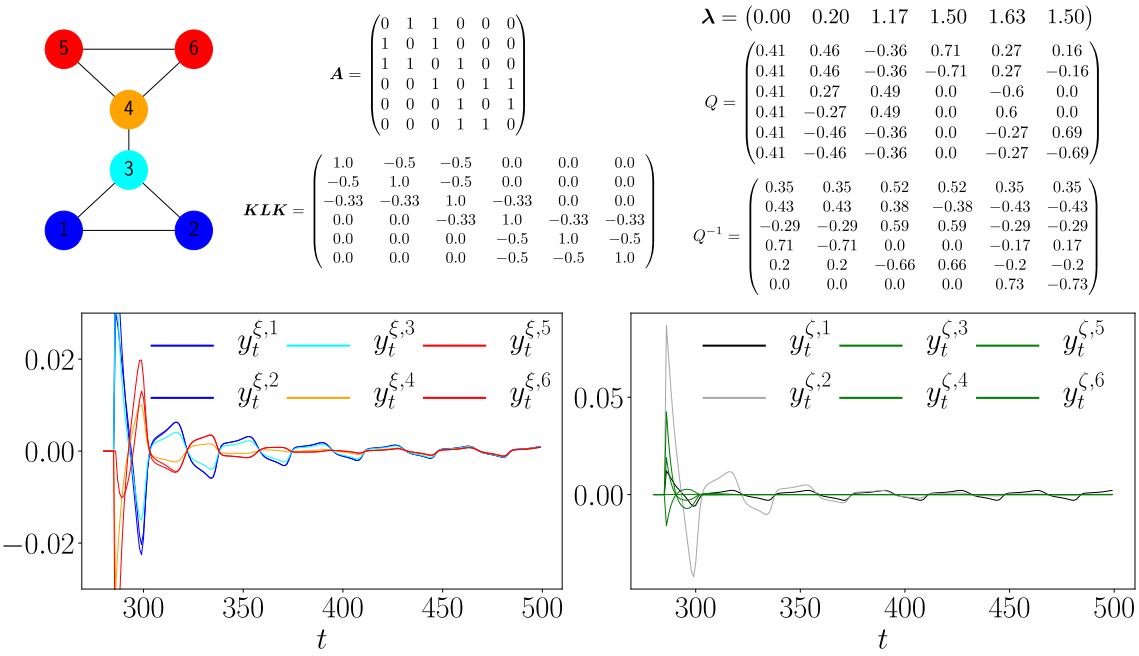


Figure 5.14: Complete synchronization in a network with six agents and two cliques of three nodes each. The top part of the figure shows the network, the adjacency matrix $\boldsymbol{A}$ and the normalized Laplacian $\boldsymbol{KLK}$. It also shows the eigenvalues $\boldsymbol{\lambda}$ and the eigenvectors $\boldsymbol{Q}$ of the normalized Laplacian, in a way that each column of $\boldsymbol{Q}$ is the eigenvector corresponding to the eigenvalue above. It finally shows the inverse matrix of eigenvectors $\boldsymbol{Q}^{-1}$. The bottom part of the figure shows the evolution of the deviations in the original basis and in the eigenbasis.

At time $\tau = 285$, we apply the following shock vector:

$$\left(y_{\tau+1}^{\xi,1}, y_{\tau+1}^{\xi,2}, y_{\tau+1}^{\xi,3}, y_{\tau+1}^{\xi,4}, y_{\tau+1}^{\xi,5}, y_{\tau+1}^{\xi,6}\right) = (0.05, 0.03, 0.04, -0.03, -0.06, 0.00). \tag{5.14}$$

This shock vector can be thought of as a positive shock hitting the clique with agents 1,2 and 3, combined with a negative shock hitting the other clique and idiosyncratic shocks hitting the agents. In the eigenbasis, this shock vector corresponds to

$$\left(y_{\tau+1}^{\zeta,1}, y_{\tau+1}^{\zeta,2}, y_{\tau+1}^{\zeta,3}, y_{\tau+1}^{\zeta,4}, y_{\tau+1}^{\zeta,5}, y_{\tau+1}^{\zeta,6}\right) = (0.01, 0.09, 0.00, 0.02, -0.02, -0.04). \tag{5.15}$$

It is already clear that the positive shocks in the first clique and negative shocks in the second clique make the initial condition of the second eigenmode largest.

The evolution of the shocks is shown at the bottom of Figure 5.14. In the panel on the left, we show the shocks in the original basis, i.e. each line corresponds to an agent. It is clear that dynamics vary across the two cliques, but they are very similar within any of the two cliques. The cyan and orange lines, corresponding to the two "bridge" agents, are the ones that are most similar to dynamics in the other clique. Over time, all dynamics converge again to the synchronized state, with a permanent phase shift.

The evolution of the shocks in the eigenbasis is shown in the right panel. The black line is the eigenmode which is parallel to the synchronization manifold. The grey line corresponds to the second eigenmode, the one with largest initial value, and decays very slowly. The green lines correspond to the third to sixth eigenvectors, and decay very quickly.

The interpretation for the eigenmodes is now clear. The second eigenmode corresponds to synchronization across cliques. Indeed, because $y_t^{\zeta,2} = 0.43y_t^{\xi,1} + 0.43y_t^{\xi,2} + 0.38y_t^{\xi,3} - 0.38y_t^{\xi,4} - 0.43y_t^{\xi,5} - 0.43y_t^{\xi,6}$, it is maximal in absolute value when shocks are positive in one clique and negative in the other one. This eigenmode takes time to decay, because the eigenvalue λ_2 is relatively small and, as shown in Figure 5.10a, the Lyapunov exponent is not very negative. This means that synchronization across cliques takes time. The third to sixth eigenmodes correspond to synchronization within cliques, and decay very quickly ($K = \epsilon\lambda_i$, $i = 3, \dots 6$ is relatively large, so the Lyapunov exponent is more negative than in the case of $K = \epsilon\lambda_2$). This indicates that synchronization within cliques happens very quickly.

This phenomenon is general, because the second eigenvector of the Laplacian—the Fiedler vector—takes positive values in a cluster of nodes, and negative values in the other. If there were M cliques, or more loosely M clusters of highly connected nodes, weakly connected across one another, there would be $M - 1$ small eigenvalues and then a large gap to the next eigenvalue. The eigenmodes corresponding to the small eigenvalues would indicate, as in the case above, synchronization across clusters. The remaining $N - M$ eigenvalues would be much larger, and the associated eigenmodes would correspond to synchronization within clusters. There could also be a hierarchy of clusters. Imagine for example that there were two main clusters, and two sub-clusters within each cluster. Here, the Fiedler vector would divide between the two main clusters, but then the following eigenmode (corresponding to a larger eigenvalue) would correspond to synchronization between sub-clusters. There would be a time-scale separation so that synchronization would first occur within sub-clusters, then across sub-clusters within the same main cluster, and finally across main clusters.

5.4.2 Input-output business cycle model

The analysis above was performed for unweighted, undirected networks. These have the convenient mathematical property that all eigenvalues of the normalized Laplacian are real, positive and bounded between zero and two. In the case of directed, weighted networks, with connectivity matrix $\boldsymbol{W}$ (so that all rows are normalized to one), instead of calculating the eigenvalues of $\boldsymbol{KLK}$ one calculates the eigenvalues of $\boldsymbol{W} - \boldsymbol{I}_N$. The analysis is similar,

but the eigenvalues could be complex.[72] In practice, with the input-output networks that we consider, the imaginary part of eigenvalues and eigenvectors is negligible, so that the analysis is very similar to the previous case.

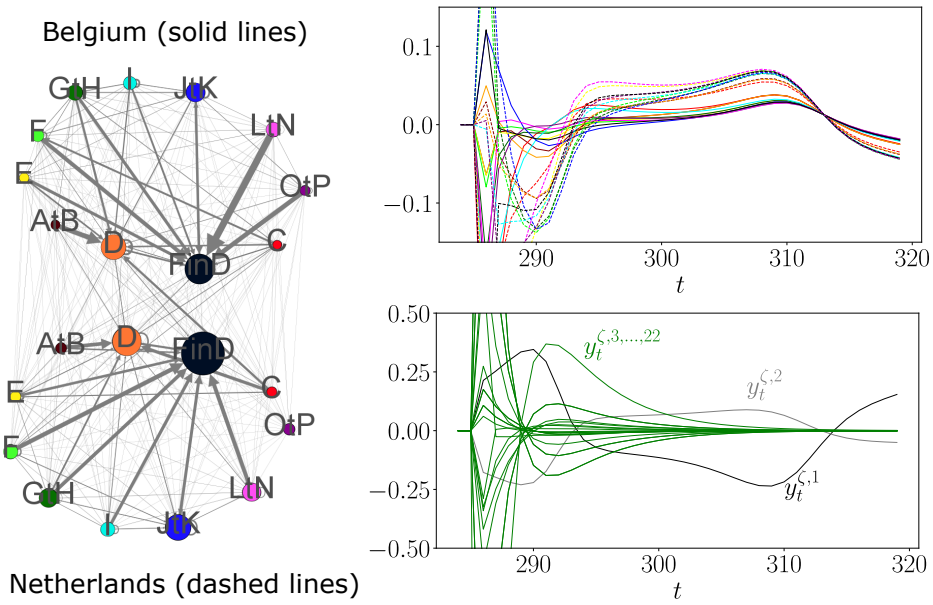


Figure 5.15: Complete synchronization in the Belgium-Netherlands input-output network. Left Panel: The node size is proportional to their total value. The links are the value flows (goods and services flow in the same direction as the arrow) and their size is proportional to the value of these flows normalized by the total output of the node. AtB: Agriculture; C: Mining; D: Manufacturing; E: Utilities; F: Construction; GtH: Trade; I: Transports; JtK: Finance and other business services; LtN: Government services; OtP: Other services; FinD: Final demand. Right panel: The lines in the top right panel are colored as the nodes in the input-output network. Solid lines correspond to Belgian sectors, while dashed lines correspond to Dutch sectors.

As an illustration, in Figure 5.15 we consider a number of mostly negative shocks hitting all economic sectors in Belgium and the Netherlands. These two countries are chosen as their output is of the same order of magnitude and, being relatively small, their openness to other countries is larger than in the case of big countries. The input-output network is shown on the left. It is composed of $N = 22$ nodes, 10 economic sectors in each country and final demand in each country. As already discussed in Section 5.3.2, the sectors that are mostly connected to the other country are AtB: Agriculture, C: Mining and D: Manufacturing. Final demand acts as a central hub within each country (the edges from final demand to the other sectors, proportional to the pointed sector's size, are only visible in the case of the "JtK: Finance and other business services" sector).

In the top right panel we show the dynamics after the shocks. The solid and dashed

[72] To apply master stability analysis in the case of complex eigenvalues is almost identical to the case of real eigenvalues, except that one computes complex Floquet exponents (instead of real Lyapunov exponents) and then verifies if the complex values of $K = \epsilon\lambda_i$ are inside the stable region.

lines quickly cluster in two distinct groups, indicating that synchronization is quickly reached within each country, but it takes longer to reach synchronization across the two countries. Interestingly, the sectors whose dynamics is closest to the other country's average dynamics are precisely agriculture, mining and manufacturing, which act as bridges between the two countries (see the dark red, red and orange lines). Since agriculture and mining are tiny as compared to manufacturing, one can say that manufacturing is the most important connection between the two countries.

In the bottom right panel we show the dynamics in the eigenbasis. The eigenvalues of $\boldsymbol{W} - \boldsymbol{I}_{22}$ are $\lambda_1 = 0$, $\lambda_2 = 0.09$, $\lambda_{3,\dots,N} \geq 0.77$.[73] The large gap between λ_2 and λ_3 indicates that the network is strongly clustered in two groups, as was evident in Figure 5.15. The Fiedler eigenvector, corresponding to λ_2, is

$$\begin{aligned}\boldsymbol{v} = (&-0.15, -0.17, -0.15, -0.25, -0.29, -0.25, -0.23, -0.26, -0.29, -0.28, -0.27,\\ &0.11, 0.04, 0.09, 0.2, 0.23, 0.19, 0.18, 0.21, 0.22, 0.21, 0.20).\end{aligned} \tag{5.16}$$

This also shows the division in two communities. The first 11 components of this vector correspond to Belgium and are negative, while the final 11 components correspond to the Netherlands and are positive. The first three components–corresponding to agriculture, mining and manufacturing–are less negative, exactly like the first three positive components which are less positive. This indicates that these sectors are closer to sectors in the other country. It also explains why the red and orange lines are close to one another in the top right panel of Figure 5.15. Because after some time only the first and second eigenmodes matter, and the first eigenmode is identical for all units, one can focus on the second eigenmode to understand the relative dynamics of the units. The deviations $y_t^{\xi,i}$ are simply proportional to the component of the Fiedler vector, and so indeed units whose components are closer in value to units in the other cluster can be expected to have closer dynamics.

Given that the spectral properties of the network are very similar to the example in Section 5.4.1, with two cliques of three nodes each, it is not surprising that dynamics are similar. The eigenmodes corresponding to eigenvalues λ_i, $i \geq 3$, decay very quickly, while the eigenmode corresponding to λ_2 decays more slowly. This is consistent with the pattern described above by which synchronization within each country is reached quickly, while synchronization across countries takes longer.

We finally consider the international input-output network in which we aggregate all sectors within a country, as we did in Section 5.3.2. This is a network with $N = 17$ nodes, whose weight matrix $\boldsymbol{W}$ is shown in Table 5.2. The eigenvalues of $\boldsymbol{W} - \boldsymbol{I}_{17}$ are: $\lambda_1 = 0.0$;

[73] Some eigenvalues have a small imaginary part. The eigenvalues with largest imaginary part are $\lambda_\pm = 0.77 \pm 0.06j$, suggesting that the real part is dominating. Simulations suggest that it should be safe to simply ignore the imaginary part.

$\lambda_2 = 0.04$; $\lambda_3 = 0.06$; $\lambda_4 = 0.08$; $\lambda_5 = 0.15$; $\lambda_6 = 0.16$; $\lambda_7 = 0.18$; $\lambda_8 = 0.2$; $\lambda_9 = 0.22$; $\lambda_{10} = 0.26$; $\lambda_{11} = 0.33$; $\lambda_{12} = 0.36$; $\lambda_{13} = 0.36$; $\lambda_{14} = 0.36$; $\lambda_{15} = 0.37$; $\lambda_{16} = 0.47$; $\lambda_{17} = 0.52$. It is clear that these eigenvalues are much smaller than the ones corresponding to the sectoral input-output network in Belgium and the Netherlands. This means that synchronization takes longer in this network, as one could expect based on how small are the international linkages relative to the national ones. Moreover, as there are no clear gaps between groups of eigenvalues, there is no clear hierarchical structure in terms of clusters or communities. This is confirmed by simulations (not reported).

It is also interesting to consider the components of the Fiedler vector. These are, with the respective country: JPN: -0.66; AUS: -0.43; CAN: 0.05; USA: 0.06; IRL: 0.1; DNK: 0.13; SWE: 0.13; FIN: 0.14; GBR: 0.14; NLD: 0.16; BEL: 0.16; DEU: 0.17; AUT: 0.17; FRA: 0.19; ITA: 0.21; PRT: 0.22; ESP: 0.23. This clearly shows geographical and cultural patterns, and it is a first approximation to which countries are likely to be most strongly correlated. (For this reason, the Fiedler vector is commonly used in spectral community detection.)

The analysis so far was performed with reference to a single shock hitting one or more agents/sector-country pairs at a given time step. Of course, in reality shocks repeatedly hit economic agents. These shocks can have a rich structure, such as heterogeneous variance, autocorrelation, and correlation structure. Here, we simply show the properties of the model with respect to repeated idiosyncratic shocks that have the same variance and autocorrelation. In the input-output interpretation, this means that we do not consider either national or sectoral shocks, and we assume that all sector-country-pair idiosyncratic shocks are homogeneous.

With shocks repeatedly hitting, the master stability approximation that we defined in Section 5.4.1 becomes worse over time, even if the shocks are small. The approximate dynamics fluctuates much more than the actual shocked dynamics, and it tends to remain in phase with the deterministic dynamics, while the phase difference between the actual shocked dynamics and the deterministic dynamics keeps changing (phase is a marginally stable variable). It is not surprising that errors keep accumulating and so the approximation becomes worse and worse. If one wanted to use the approximation over long time spans, it should be necessary to periodically realign the approximation to the actual shocked dynamics rather than to the deterministic one.[74]

Here, we simply reconsider the example of 10 sectors and final demand in Belgium and the Netherlands that was introduced above. We let each node (sector or final demand in

[74] This would be similar to the periodic orthonormalization that is necessary for the calculation of Lyapunov exponents (Benettin et al., 1980).

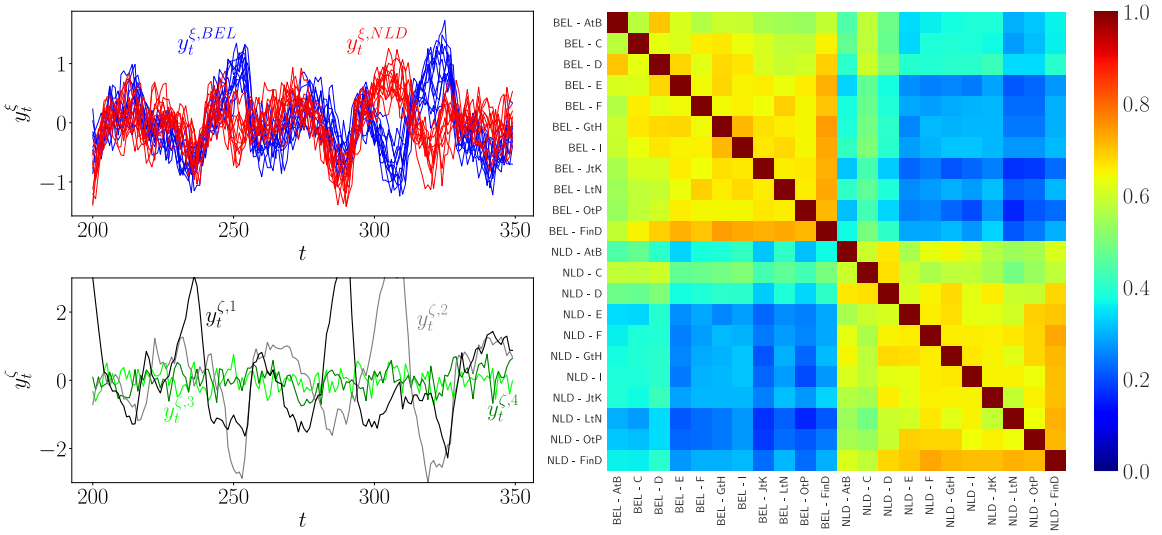


Figure 5.16: Complete synchronization under repeated shocks in the Belgium-Netherlands input-output network. Left panels: dynamics of the deviations in the original basis and in the eigenbasis. Right panel: correlation matrix between sectors and final demand in the two countries.

one of the two countries) be hit each time step by an AR(1) shock process with correlation $\rho_u = 0.3$ and normal innovations with zero mean and standard deviation $\sigma_u = 0.2$. The shocks are relatively large, leading to very irregular dynamics of the units which, however, retain a certain periodicity. We are imposing such strong shocks to see how predictive is the theory in this extreme benchmark.

In Figure 5.16, top left panel, we show the deviations from the limit cycle $y_t^{\xi,i}$.[75] We color blue the deviations of Belgian sectors, and red those of Dutch sectors. It is clear that deviations tend to be more similar within than across countries, although we imposed no country-specific shock. In the bottom left panel we show the dynamics of the first four eigenmodes. The variance of the first two eigenmodes is much larger than that of the subsequent eigenmodes. This is due to the fact that the master stability function suppresses the eigenmodes with larger eigenvalues than λ_2 much more quickly. When the dynamics become very different across the two countries, the second eigenmode is large, while, when the dynamics are similar across the two countries, it is small.

On the right, we show a correlation matrix, obtained by averaging the actual dynamics (not the approximated one) over 1000 time steps. It is clear that sectors are much more correlated within countries than across countries. It is also clear that the most internationally correlated sectors are manufacturing, agriculture and mining, followed by "finance and

[75] As discussed above, the sum of the unperturbed deterministic limit cycles and these deviations calculated using Eq. (5.10) does not generally correspond to the actual perturbed dynamics. However, the properties that we will discuss below are identical across the actual dynamics and the sum of the deterministic limit cycle and the deviations.

other business" and final demand. The least correlated sectors are government and other services. Within countries, final demand is highly correlated to all sectors. Agriculture and mining are the least correlated sectors.

The important question that we discussed at the beginning of this section on complete synchronization, Section 5.4, is whether the combination of synchronizing non-linear dynamics and shock propagation could generate stronger comovement across sectors and countries than shock propagation only (Eqs. 5.7 and 5.8). When shocks are small, correlations among time series are mostly driven by the synchronization of the underlying non-linear dynamics while, when shocks are large, correlations are driven by the comovement of the shocks. Preliminary simulations seem to suggest that the co-existence of underlying non-linear dynamics and shocks gives higher correlations than the simple propagation of shocks if shocks are not too large (relative to the amplitude of limit cycles). Overall, this confirms the promise of synchronization models to account for the relatively high degree of comovement across sectors and countries, explaining at once how national business cycle emerge from individual fluctuations and how international business cycles arise.

5.5 Extensions

The work presented so far is somewhat preliminary, in the sense that, while it illustrates the mathematical mechanisms that lead to synchronization, it lacks a quantitative application that explicitly targets properties of the data. Although the model presented here could be calibrated with that goal, modifications of the model are likely to make it score better. In the following, we give a list of extensions that have this potential.

Microfoundations. The dynamical equations of the model presented in Section 4.3.2 are directly assumed, rather than being derived from explicit assumptions about the behavior of agents. In fact, Beaudry et al. (2019) show how these dynamical equations could be obtained from microfoundations, staying within the framework of Dynamic Stochastic General Equilibrium (DSGE) models featuring rational expectations and intertemporal optimization. It is not clear to us if the DSGE framework can deal with the strong heterogeneity across different agents that we need to assume to study synchronization. An alternative could be to microfound the dynamical equations in Eq. (4.2) using an Agent-Based Model. Either way, an advantage of having a microfounded model is a clearer interpretation of the variables and parameters. The main disadvantage is that it makes the synchronization techniques look specific to that problem. Ideally, one could use the general framework based on strategic complementarities to illustrate synchronization methods, and then perform the quantitative analysis with a microfounded model.

Heterogeneity. Although in principle the basic units of our model are agents (of any type), in the practical applications we considered sector-country pairs as the basic unit of analysis. Under this assumption, companies in an entire sector in a given country behave in the same way. This assumption is known as the *representative agent* in macroeconomics. Our choice was dictated by data availability—we used the sectoral input-output network as the agents' interaction structure. Yet, with more disaggregated data becoming available, one could consider smaller aggregates. For example, data on firm-to-firm production networks should make it possible to calibrate the model in a way that agents correspond to individual firms. It is also possible to disaggregate the final demand node and consider heterogeneous households. Each household may have different tastes for goods, splitting up final demand to each sector-country pair in a heterogeneous way. Households may also be connected among themselves in a social network, for example copying consumption patterns of their friends (Asano et al., 2019). Considering heterogeneity at the household level would allow studying inequality dynamics.

Evolving input-output network. In our analysis, we assumed the input-output network to be constant. This is not too of a strong assumption in some cases, as quantitatively the input-output network changes slowly (Cerina et al., 2015). Better quantitative predictions can however certainly be obtained by including the evolving nature of the input-output network in the analysis. There may also be feedback loops between the evolution of the interaction network and the shorter-term business cycle dynamics. For example, asymmetric business cycle dynamics could induce different countries in specializing in different sectors.

Other synchronization channels. Our analysis includes three out of six of the main synchronization channels surveyed in Section 4.4.1. It includes trade, common fiscal and monetary policy (country- or monetary union-specific shocks) and similar industrial structure (sector-specific shocks). It does not include finance, multinational corporate control, and diffusion of animal spirits. While the role of finance is debated, at least during the 2008 financial crisis it seemed to have a significant role. Finance could be included in the aggregate, by using data on bilateral financial exposures across countries (Lane and Milesi-Ferretti, 2007), or at the individual level, by augmenting the model with bank agents. Combining the trade and financial networks could be best modeled within the context of multilayer networks (Kivelä et al., 2014). An interesting recent work that studied propagation of exogenous shocks on a trade-finance multilayer network is Starnini et al. (2019). Differently from finance, the synchronizing effect of multinational corporate control is uncontroversial, and probably quantitatively relevant (Di Giovanni et al., 2018). By considering individual

firms as the basic unit of analysis, we could certainly extend our synchronization framework to include this channel. Finally, by including expectations we could include propagation of animal spirits. Under this channel, for example, agents in a given country could read on the newspaper that a closely related country is on the point of entering a recession, become pessimistic, and trigger a recession within the country of origin (self-fulfilling beliefs).

Chaos. Non-linear dynamics in our model corresponds to limit cycles. This requires augmenting the model with relatively strong exogenous shocks to make time series generated by the model resemble empirical one, which do not show clear periodicity. A complementary approach would be to consider a chaotic business cycle model, like some of the ones surveyed in Section 4.3.1. A clear advantage is that one would require much weaker shocks—or even no shocks at all—to reproduce the periodicity of empirical time series. The main difficulty with this approach is that there exist no general results about sufficient conditions for chaos in relatively high-dimensional systems, and so it is not clear a priori which model specifications could be seen as promising. Moreover, many chaotic attractors look almost periodic at least for certain periods, removing the main advantage of using a chaotic model.

Empirical validation of the shocks. In our analysis, we have considered rather abstract idiosyncratic and common shocks. To build a quantitative model, one would need to define shock processes in line with the data. Finding shocks directly from the data has historically been impossible in all but a handful of cases; to identify shocks, researchers need to make strong assumptions about the statistical properties of the data. Some recent work that uses Natural Language Processing methods to identify shocks from newspaper articles (Thorsrud, 2018) seems promising in this direction.

Regional business cycles. One could also consider regional aggregates in place of sectoral aggregates. This would allow linking this work to research on resilience in economic geography (Rocchetta and Mina, 2019). It would also make it possible to study the synchronization of cycles in housing market prices across regions.

This long list of extensions highlights the potential of synchronization theory to address many issues in economics. Economic modelers only need to assume that fluctuations are at least in part explained by endogenous non-linear dynamics: once these dynamics are assumed, synchronization in the sense presented in this thesis becomes a central element.

Conclusions

In this thesis, we have provided theoretical arguments to why fluctuations at the smallest and largest economic scales may arise as an endogenous phenomenon. When two players repeatedly play a game, their learning dynamics may converge to a Nash equilibrium or may endogenously fluctuate. We have shown that equilibrium convergence becomes less and less likely as games become more complicated and more competitive, as in many situations in the real world. At the largest scale of macroeconomic systems, we have introduced a mathematical framework to explain synchronization of endogenous business cycles. We have argued that synchronization of non-linear dynamics is a channel to generate relatively high comovement across economic sectors and countries. Given the difficulty of models based on exogenous shocks to generate a level of comovement that is as high as in the data, we view this property as a potentially compelling indirect evidence of endogenous fluctuations in macroeconomics.

In Sections 3.7 and 5.5, we have discussed perspectives for our work on games and macroeconomics, respectively. While many theoretical extensions are possible, we are particularly interested in taking our results to data. This is somewhat difficult in the case of games, because it is generally not possible to directly observe learning dynamics of players. However, carefully designed experiments such as Cason et al. (2013) should make it possible to experimentally test the equilibrium convergence theory presented in this thesis. It should be more straightforward to quantitatively check the potential of endogenous business cycle models to generate comovement: We plan to work on that in the forthcoming months.

We think of the economy as a complex system, formed by billions of heterogeneous agents that learn and interact, producing aggregate phenomena that do not directly match their individual behaviors. We view the emergence of endogenous fluctuations at several scales as a distinctive signature of the complexity of the economy.

Bibliography

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012) "The network origins of aggregate fluctuations," *Econometrica*, Vol. 80, pp. 1977–2016.

Amiet, B., Collevecchio, A., and Scarsini, M. (2019) "Pure Nash equilibria and best-response dynamics in random games," *arXiv preprint arXiv:1905.10758*.

Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y., and Zhou, C. (2008) "Synchronization in complex networks," *Physics Reports*, Vol. 469, pp. 93–153.

Arieli, I. and Young, H. P. (2016) "Stochastic learning dynamics and speed of convergence in population games," *Econometrica*, Vol. 84, pp. 627–676.

Arkolakis, C. and Ramanarayanan, A. (2009) "Vertical specialization and international business cycle synchronization," *Scandinavian Journal of Economics*, Vol. 111, pp. 655–680.

Arrow, K. J. and Hurwicz, L. (1960) "Stability of the gradient process in n-person games," *Journal of the Society for Industrial and Applied Mathematics*, Vol. 8, pp. 280–294.

Asano, Y. M., Kolb, J. J., Heitzig, J., and Farmer, J. D. (2019) "Emergent inequality and endogenous dynamics in a simple behavioral macroeconomic model," *arXiv preprint arXiv:1907.02155*.

Aumann, R. J. (1987) "Correlated equilibrium as an expression of Bayesian rationality," *Econometrica*, pp. 1–18.

Azariadis, C. and Smith, B. (1998) "Financial intermediation and regime switching in business cycles," *American Economic Review*, pp. 516–536.

Backus, D. K., Kehoe, P. J., and Kydland, F. E. (1992) "International real business cycles," *Journal of Political Economy*, Vol. 100, pp. 745–775.

Baqaee, D. R. and Farhi, E. (2019) "The macroeconomic impact of microeconomic shocks: beyond Hulten's Theorem," *Econometrica*, Vol. 87, pp. 1155–1203.

Barahona, M. and Pecora, L. M. (2002) "Synchronization in small-world systems," *Physical Review Letters*, Vol. 89, p. 054101.

Barnett, W. A. and Chen, P. (1988) "The aggregation-theoretic monetary aggregates are chaotic and have strange attractors," in W. A. Barnett, E. R. Berndt, and H. White eds. *Dynamic Econometric Modeling*: Cambridge University Press, pp. 199–246.

Barnett, W. A. and Serletis, A. (2000) "Martingales, nonlinearity, and chaos," *Journal of Economic Dynamics and Control*, Vol. 24, pp. 703–724.

Baxter, M. and Kouparitsas, M. A. (2005) "Determinants of business cycle comovement: a robust analysis," *Journal of Monetary Economics*, Vol. 52, pp. 113–157.

Beaudry, P., Galizia, D., and Portier, F. (2015) "Reviving the Limit Cycle View of Macroeconomic Fluctuations,"Technical report, National Bureau of Economic Research.

——— (2019) "Putting the Cycle Back into Business Cycle Analysis," *Forthcoming, American Economic Review*.

Benaïm, M. and Hirsch, M. W. (1999) "Mixed equilibria and dynamical systems arising from fictitious play in perturbed games," *Games and Economic Behavior*, Vol. 29, pp. 36–72.

Benaïm, M., Hofbauer, J., and Hopkins, E. (2009) "Learning in games with unstable equilibria," *Journal of Economic Theory*, Vol. 144, pp. 1694–1709.

Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J.-M. (1980) "Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory," *Meccanica*, Vol. 15, pp. 9–20.

Benhabib, J. and Day, R. H. (1981) "Rational choice and erratic behaviour," *Review of Economic Studies*, Vol. 48, pp. 459–471.

——— (1982) "A characterization of erratic dynamics in the overlapping generations model," *Journal of Economic Dynamics and Control*, Vol. 4, pp. 37–55.

Benhabib, J. and Farmer, R. E. (1999) "Indeterminacy and sunspots in macroeconomics," *Handbook of Macroeconomics*, Vol. 1, pp. 387–448.

Benhabib, J. and Nishimura, K. (1979) "The Hopf bifurcation and existence and stability of closed orbits in multisector models of optimal economic growth," *Journal of Economic Theory*, Vol. 21, pp. 421–444.

Benhabib, J., Rogerson, R., and Wright, R. (1991) "Homework in macroeconomics: Household production and aggregate fluctuations," *Journal of Political Economy*, Vol. 99, pp. 1166–1187.

Berg, J. and Engel, A. (1998) "Matrix games, mixed strategies, and statistical mechanics," *Physical Review Letters*, Vol. 81, pp. 4999–5002.

Berg, J. (2000) "Statistical mechanics of random two-player games," *Physical Review E*, Vol. 61, pp. 2327–2339.

Berg, J. and Weigt, M. (1999) "Entropy and typical properties of Nash equilibria in two-player games," *EPL (Europhysics Letters)*, Vol. 48, pp. 129–135.

Bischi, G. I., Chiarella, C., Kopel, M., and Szidarovszky, F. (2009) *Nonlinear oligopolies: Stability and bifurcations*: Springer Science & Business Media.

Bloomfield, R. (1994) "Learning a mixed strategy equilibrium in the laboratory," *Journal of Economic Behavior & Organization*, Vol. 25, pp. 411–436.

Blum, A. and Mansour, Y. (2007) "Learning, regret minimization, and equilibria," in N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani eds. *Algorithmic Game Theory*: Cambridge University Press, pp. 79–102.

Blume, L. E. and Easley, D. (1982) "Learning to be rational," *Journal of Economic Theory*, Vol. 26, pp. 340–351.

Boldrin, M. and Montrucchio, L. (1986) "On the indeterminacy of capital accumulation paths," *Journal of Economic Theory*, Vol. 40, pp. 26–39.

Boldrin, M. and Woodford, M. (1990) "Equilibrium models displaying endogenous fluctuations and chaos: a survey," *Journal of Monetary Economics*, Vol. 25, pp. 189–222.

Bonart, J., Bouchaud, J.-P., Landier, A., and Thesmar, D. (2014) "Instabilities in large economies: aggregate volatility without idiosyncratic shocks," *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2014, p. P10040.

Börgers, T. and Sarin, R. (1997) "Learning through reinforcement and replicator dynamics," *Journal of Economic Theory*, Vol. 77, pp. 1–14.

Brock, W. A. (1986) "Distinguishing random and deterministic systems: Abridged version," *Journal of Economic Theory*, Vol. 40, pp. 168–195.

Brock, W. A. and Sayers, C. L. (1988) "Is the business cycle characterized by deterministic chaos?" *Journal of Monetary Economics*, Vol. 22, pp. 71–90.

Brown, G. W. (1951) "Iterative solution of games by fictitious play," in T. Koopmans ed. *Activity analysis of production and allocation*, New York: Wiley, pp. 374–376.

Bullard, J. (1994) "Learning equilibria," *Journal of Economic Theory*, Vol. 64, pp. 468–485.

Bulow, J. I., Geanakoplos, J. D., and Klemperer, P. D. (1985) "Multimarket oligopoly: Strategic substitutes and complements," *Journal of Political Economy*, Vol. 93, pp. 488–511.

Burns, A. F. and Mitchell, W. C. (1946) "Measuring business cycles," *NBER Books*.

Bush, R. R. and Mosteller, F. (1955) *Stochastic models for learning.*: John Wiley & Sons, Inc.

Calderon, C., Chong, A., and Stein, E. (2007) "Trade intensity and business cycle synchronization: Are developing countries any different?" *Journal of International Economics*, Vol. 71, pp. 2–21.

Camerer, C. (2003) *Behavioral game theory: Experiments in strategic interaction*: Princeton University Press.

Camerer, C. F., Ho, T.-H., and Chong, J.-K. (2002) "Sophisticated experience-weighted attraction learning and strategic teaching in repeated games," *Journal of Economic Theory*, Vol. 104, pp. 137–188.

Camerer, C. and Ho, T. (1999) "Experience-weighted attraction learning in normal form games," *Econometrica*, Vol. 67, pp. 827–874.

Candogan, O., Ozdaglar, A., and Parrilo, P. A. (2013) "Dynamics in near-potential games," *Games and Economic Behavior*, Vol. 82, pp. 66–90.

Carvalho, V. M. (2010) "Aggregate fluctuations and the network structure of intersectoral trade," *Working Paper, CREI*.

Cason, T. N., Friedman, D., and Hopkins, E. (2013) "Cycles and instability in a rock–paper–scissors population game: a continuous time experiment," *Review of Economic Studies*, Vol. 81, pp. 112–136.

Cerina, F., Zhu, Z., Chessa, A., and Riccaboni, M. (2015) "World input-output network," *PloS One*, Vol. 10, p. e0134025.

Chawanya, T. (1995) "A new type of irregular motion in a class of game dynamics systems," *Progress of Theoretical Physics*, Vol. 94, pp. 163–179.

Cheung, Y.-W. and Friedman, D. (1997) "Individual learning in normal form games: Some laboratory results," *Games and Economic Behavior*, Vol. 19, pp. 46–76.

Christiano, L. J. and Fitzgerald, T. J. (1998) "The Business Cycle: It's Still a Puzzle," *Economic Perspectives*, Vol. 22, pp. 56–83.

——— (2003) "The band pass filter," *International Economic Review*, Vol. 44, pp. 435–465.

Clark, T. E. and Van Wincoop, E. (2001) "Borders and business cycles," *Journal of International Economics*, Vol. 55, pp. 59–85.

Cohen, J. E. (1998) "Cooperation and self-interest: Pareto-inefficiency of Nash equilibria in finite random games," *Proceedings of the National Academy of Sciences*, Vol. 95, pp. 9724–9731.

Conlisk, J. (1993a) "Adaptation in games: Two solutions to the Crawford puzzle," *Journal of Economic Behavior & Organization*, Vol. 22, pp. 25–50.

——— (1993b) "Adaptive tactics in games: Further solutions to the Crawford puzzle," *Journal of Economic Behavior & Organization*, Vol. 22, pp. 51–68.

Cooper, R. and Haltiwanger, J. (1990) "Inventories and the propagation of sectoral shocks," *American Economic Review*, pp. 170–190.

Cooper, R. and John, A. (1988) "Coordinating coordination failures in Keynesian models," *Quarterly Journal of Economics*, pp. 441–463.

Cournot, A.-A. (1838) *Recherches sur les principes mathématiques de la théorie des richesses*: L. Hachette.

Cravino, J. and Levchenko, A. A. (2017) "Multinational firms and international business cycle transmission," *Quarterly Journal of Economics*, Vol. 132, pp. 921–962.

Crawford, V. P. (1974) "Learning the optimal strategy in a zero-sum game," *Econometrica*, pp. 885–891.

——— (1985) "Learning behavior and mixed-strategy Nash equilibria," *Journal of Economic Behavior & Organization*, Vol. 6, pp. 69–78.

Crawford, V. P., Costa-Gomes, M. A., and Iriberri, N. (2013) "Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications," *Journal of Economic Literature*, Vol. 51, pp. 5–62.

Cross, J. G. (1973) "A stochastic learning model of economic behavior," *Quarterly Journal of Economics*, Vol. 87, pp. 239–266.

Crutchfield, J. P., Farmer, J. D., and Huberman, B. A. (1982) "Fluctuations and simple chaotic dynamics," *Physics Reports*, Vol. 92, pp. 45–82.

Dawid, H. and Delli Gatti, D. (2018) "Agent-based macroeconomics," in *Handbook of Computational Economics*, Vol. 4: Elsevier, pp. 63–156.

Day, R. H. (1982) "Irregular growth cycles," *American Economic Review*, Vol. 72, pp. 406–414.

De Grauwe, P. (2011) "Animal spirits and monetary policy," *Economic Theory*, Vol. 47, pp. 423–457.

De Haan, J., Inklaar, R., and Jong-A-Pin, R. (2008) "Will business cycles in the euro area converge? A critical survey of empirical research," *Journal of Economic Surveys*, Vol. 22, pp. 234–273.

DeCoster, G. P. and Mitchell, D. W. (1991) "Nonlinear monetary dynamics," *Journal of Business & Economic Statistics*, Vol. 9, pp. 455–461.

Desai, M. A. and Foley, C. F. (2004) "The comovement of returns and investment within the multinational firm,"Technical report, National Bureau of Economic Research.

Devereux, M. B. and Yetman, J. (2010) "Leverage constraints and the international transmission of shocks," *Journal of Money, Credit and Banking*, Vol. 42, pp. 71–105.

Di Giovanni, J. and Levchenko, A. A. (2010) "Putting the parts together: trade, vertical linkages, and business cycle comovement," *American Economic Journal: Macroeconomics*, Vol. 2, pp. 95–124.

Di Giovanni, J., Levchenko, A. A., and Mejean, I. (2018) "The micro origins of international business-cycle comovement," *American Economic Review*, Vol. 108, pp. 82–108.

Diamond, P. and Fudenberg, D. (1989) "Rational expectations business cycles in search equilibrium," *Journal of Political Economy*, Vol. 97, pp. 606–619.

Diederich, S. and Opper, M. (1989) "Replicators with random interactions: A solvable model," *Physical Review A*, Vol. 39, pp. 4333–4337.

Dresher, M. (1970) "Probability of a pure equilibrium point in n-person games," *Journal of Combinatorial Theory*, Vol. 8, pp. 134–145.

Dupor, B. (1999) "Aggregation and irrelevance in multi-sector models," *Journal of Monetary Economics*, Vol. 43, pp. 391–409.

Easley, D. and Kleinberg, J. (2010) *Networks, crowds, and markets: Reasoning about a highly connected world*: Cambridge University Press.

Eckmann, J.-P., Kamphorst, S. O., Ruelle, D., and Ciliberto, S. (1986) "Liapunov exponents from time series," *Physical Review A*, Vol. 34, p. 4971.

Englund, P., Persson, T., and Svensson, L. E. (1992) "Swedish business cycles: 1861–1988," *Journal of Monetary Economics*, Vol. 30, pp. 343–371.

Erev, I. and Roth, A. E. (1998) "Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria," *American Economic Review*, Vol. 88, pp. 848–881.

Erev, I., Roth, A. E., Slonim, R., and Barron, G. (2002) "Predictive value and the usefulness of game theoretic models," *International Journal of Forecasting*, Vol. 18, pp. 359–368.

Erev, I., Roth, A. E., Slonim, R. L., and Barron, G. (2007) "Learning and equilibrium as useful approximations: Accuracy of prediction on randomly selected constant sum games," *Economic Theory*, Vol. 33, pp. 29–51.

Evans, G. W. and Honkapohja, S. (2012) *Learning and expectations in macroeconomics*: Princeton University Press.

Evans, T. (2013) "K-level reasoning: A dynamic model of game learning," Master thesis, University of Manchester.

Farmer, J. D., Ott, E., and Yorke, J. A. (1983) "The dimension of chaotic attractors," *Physica D: Nonlinear Phenomena*, Vol. 7, pp. 153–180.

Foerster, A. T., Sarte, P.-D. G., and Watson, M. W. (2011) "Sectoral versus aggregate shocks: A structural factor analysis of industrial production," *Journal of Political Economy*, Vol. 119, pp. 1–38.

Foley, D. K. (1992) "A contribution to the theory of business cycles," *Quarterly Journal of Economics*, Vol. 107, pp. 1071–1088.

Forni, M. and Reichlin, L. (1998) "Let's get real: a factor analytical approach to disaggregated business cycle dynamics," *Review of Economic Studies*, Vol. 65, pp. 453–473.

Foster, D. P. and Vohra, R. V. (1997) "Calibrated learning and correlated equilibrium," *Games and Economic Behavior*, Vol. 21, pp. 40–55.

Foster, D. P. and Young, H. P. (1998) "On the nonconvergence of fictitious play in coordination games," *Games and Economic Behavior*, Vol. 25, pp. 79–96.

——— (2001) "On the impossibility of predicting the behavior of rational agents," *Proceedings of the National Academy of Sciences*, Vol. 98, pp. 12848–12853.

——— (2003) "Learning, hypothesis testing, and Nash equilibrium," *Games and Economic Behavior*, Vol. 45, pp. 73–96.

——— (2006) "Regret testing: Learning to play Nash equilibrium without knowing you have an opponent," *Theoretical Economics*, Vol. 1, pp. 341–367.

Frank, M., Gencay, R., and Stengos, T. (1988) "International chaos?" *European Economic Review*, Vol. 32, pp. 1569–1584.

Frank, M. Z. and Stengos, T. (1988) "Some evidence concerning macroeconomic chaos," *Journal of Monetary Economics*, Vol. 22, pp. 423–438.

Frankel, J. A. and Rose, A. K. (1998) "The endogenity of the optimum currency area criteria," *Economic Journal*, Vol. 108, pp. 1009–1025.

Frey, S. and Goldstone, R. L. (2013) "Cyclic game dynamics driven by iterated reasoning," *PloS One*, Vol. 8, p. e56416.

Friedman, D. (1996) "Equilibrium in evolutionary games: Some experimental results," *Economic Journal*, pp. 1–25.

Frisch, R. (1933) "Propagation problems and impulse problems in dynamic economics," in J. Taylor and M. Woodford eds. *Economic Essays in Honour of Gustav Cassel*: Allen & Unwin, London.

Fudenberg, D. and Kreps, D. M. (1993) "Learning mixed equilibria," *Games and Economic Behavior*, Vol. 5, pp. 320–367.

Fudenberg, D. and Levine, D. K. (1998) *The theory of learning in games*, Vol. 2: MIT press.

Galizia, D. (2018) "Saddle Cycles: Solving Rational Expectations Models Featuring Limit Cycles (or Chaos) Using Perturbation Methods,"Technical report.

Galla, T. (2009) "Intrinsic noise in game dynamical learning," *Physical Review Letters*, Vol. 103, p. 198702.

Galla, T. and Farmer, J. D. (2013) "Complex dynamics in learning complicated games," *Proceedings of the National Academy of Sciences*, Vol. 110, pp. 1232–1236.

Gandolfo, G. (2009) *Economic dynamics: Fourth edition*: Springer Verlag.

Germano, F. and Lugosi, G. (2007) "Global Nash convergence of Foster and Young's regret testing," *Games and Economic Behavior*, Vol. 60, pp. 135–154.

Gigerenzer, G. and Todd, P. M. (1999) *Simple heuristics that make us smart*: Oxford University Press.

Gintis, H. (2000) *Game theory evolving: A problem-centered introduction to modeling strategic behavior*: Princeton University Press.

Goemans, M., Mirrokni, V., and Vetta, A. (2005) "Sink equilibria and convergence," in *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*, pp. 142–151, IEEE.

Gokhale, C. S. and Traulsen, A. (2010) "Evolutionary games in the multiverse," *Proceedings of the National Academy of Sciences*, Vol. 107, pp. 5500–5504.

Goldberg, K., Goldman, A., and Newman, M. (1968) "The probability of an equilibrium point," *Journal of Research of the National Bureau of Standards*, Vol. 72, pp. 93–101.

Gómez-Gardenes, J., Moreno, Y., and Arenas, A. (2007) "Paths to synchronization on complex networks," *Physical Review Letters*, Vol. 98, p. 034101.

Goodwin, R. M. (1951) "The nonlinear accelerator and the persistence of business cycles," *Econometrica*, pp. 1–17.

——— (1967) *A Growth Cycle*: Cambridge University Press.

Grandmont, J.-M. (1985) "On endogenous competitive business cycles," *Econometrica*, Vol. 53, pp. 995–1045.

Granger, C. W. (1966) "The typical spectral shape of an economic variable," *Econometrica*, pp. 150–161.

Grassberger, P. and Procaccia, I. (1983) "Measuring the strangeness of strange attractors," *Physica D: Nonlinear Phenomena*, Vol. 9, pp. 189–208.

Gualdi, S., Bouchaud, J.-P., Cencetti, G., Tarzia, M., and Zamponi, F. (2015) "Endogenous crisis waves: stochastic model with synchronized collective behavior," *Physical Review Letters*, Vol. 114, p. 088701.

Hahn, S. (1999) "The convergence of fictitious play in 3×3 games with strategic complementarities," *Economics Letters*, Vol. 64, pp. 57–60.

Han, T. A., Traulsen, A., and Gokhale, C. S. (2012) "On equilibrium properties of evolutionary multi-player games with random payoff matrices," *Theoretical Population Biology*, Vol. 81, pp. 264–272.

Hart, S. (2005) "Adaptive heuristics," *Econometrica*, Vol. 73, pp. 1401–1430.

Hart, S. and Mas-Colell, A. (2000) "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, Vol. 68, pp. 1127–1150.

Heathcote, J. and Perri, F. (2004) "Financial globalization and real regionalization," *Journal of Economic Theory*, Vol. 1, pp. 207–243.

Helbing, D., Lämmer, S., Witt, U., and Brenner, T. (2004) "Network-induced oscillatory behavior in material flow networks and irregular business cycles," *Physical Review E*, Vol. 70, p. 056118.

Hicks, J. R. (1950) *A Contribution to the Theory of the Trade Cycle*: Oxford University Press.

Ho, T. H., Camerer, C. F., and Chong, J.-K. (2007) "Self-tuning experience weighted attraction learning in games," *Journal of Economic Theory*, Vol. 133, pp. 177–198.

Hofbauer, J. and Sandholm, W. H. (2002) "On the global convergence of stochastic fictitious play," *Econometrica*, Vol. 70, pp. 2265–2294.

Hofbauer, J. and Sigmund, K. (1998) *Evolutionary games and population dynamics*: Cambridge university press.

Hofbauer, J. and Sorin, S. (2006) "Best response dynamics for continuous zero-sum games," *Discrete and Continuous Dynamical Systems Series B*, Vol. 6, p. 215.

Hofbauer, J. and Weibull, J. W. (1996) "Evolutionary selection against dominated strategies," *Journal of Economic Theory*, Vol. 71, pp. 558–573.

Hommes, C. H. and Manzan, S. (2006) "Comments on "Testing for nonlinear structure and chaos in economic time series"," *Journal of Macroeconomics*, Vol. 28, pp. 169–174.

Hommes, C. and Sorger, G. (1998) "Consistent expectations equilibria," *Macroeconomic Dynamics*, Vol. 2, pp. 287–321.

Hopkins, E. (1999) "A note on best response dynamics," *Games and Economic Behavior*, Vol. 29, pp. 138–150.

——— (2002) "Two competing models of how people learn in games," *Econometrica*, Vol. 70, pp. 2141–2166.

Horii, R. (2012) "Wants and past knowledge: Growth cycles with emerging industries," *Journal of Economic Dynamics and Control*, Vol. 36, pp. 220–238.

Hornstein, A. and Praschnik, J. (1997) "Intermediate inputs and sectoral comovement in the business cycle," *Journal of Monetary Economics*, Vol. 40, pp. 573–595.

Horvath, M. (1998) "Cyclicality and sectoral linkages: Aggregate fluctuations from independent sectoral shocks," *Review of Economic Dynamics*, Vol. 1, pp. 781–808.

——— (2000) "Sectoral shocks and aggregate fluctuations," *Journal of Monetary Economics*, Vol. 45, pp. 69–106.

Hsu, C.-C., Wu, J.-Y., and Yau, R. (2011) "Foreign direct investment and business cycle co-movements: The panel data evidence," *Journal of Macroeconomics*, Vol. 33, pp. 770–783.

Huberman, B. A. and Hogg, T. (1988) "The behavior of computational ecologies," in *The Ecology of Computation*, pp. 77–115: North-Holland.

Huffman, G. W. and Wynne, M. A. (1999) "The role of intratemporal adjustment costs in a multisector economy," *Journal of Monetary Economics*, Vol. 43, pp. 317–350.

Imbs, J. (2004) "Trade, finance, specialization, and synchronization," *Review of Economics and Statistics*, Vol. 86, pp. 723–734.

——— (2006) "The real effects of financial integration," *Journal of International Economics*, Vol. 68, pp. 296–324.

Inklaar, R., Jong-A-Pin, R., and De Haan, J. (2008) "Trade and business cycle synchronization in OECD countries—A re-examination," *European Economic Review*, Vol. 52, pp. 646–666.

Johnson, R. C. (2014) "Trade in intermediate inputs and business cycle comovement," *American Economic Journal: Macroeconomics*, Vol. 6, pp. 39–83.

Johnson, S., Domínguez-García, V., Donetti, L., and Muñoz, M. A. (2014) "Trophic coherence determines food-web stability," *Proceedings of the National Academy of Sciences*, Vol. 111, pp. 17923–17928.

Jordan, J. S. (1991) "Bayesian learning in normal form games," *Games and Economic Behavior*, Vol. 3, pp. 60–81.

——— (1993) "Three problems in learning mixed-strategy Nash equilibria," *Games and Economic Behavior*, Vol. 5, pp. 368–386.

Judd, K. L. (1985) "On the performance of patents," *Econometrica*, pp. 567–585.

Kalai, E. and Lehrer, E. (1993) "Rational learning leads to Nash equilibrium," *Econometrica*, pp. 1019–1045.

Kaldor, N. (1940) "A model of the trade cycle," *Economic Journal*, pp. 78–92.

Kalemli-Ozcan, S., Papaioannou, E., and Peydro, J.-L. (2013) "Financial regulation, financial globalization, and the synchronization of economic activity," *Journal of Finance*, Vol. 68, pp. 1179–1228.

Kalemli-Ozcan, S., Sørensen, B. E., and Yosha, O. (2001) "Economic integration, industrial specialization, and the asymmetry of macroeconomic fluctuations," *Journal of International Economics*, Vol. 55, pp. 107–137.

Keen, S. (1995) "Finance and economic breakdown: modeling Minsky's financial instability hypothesis," *Journal of Post Keynesian Economics*, Vol. 17, pp. 607–635.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014) "Multilayer networks," *Journal of Complex Networks*, Vol. 2, pp. 203–271.

Kleinert, J., Martin, J., and Toubal, F. (2015) "The few leading the many: Foreign affiliates and business cycle comovement," *American Economic Journal: Macroeconomics*, Vol. 7, pp. 134–59.

Kose, M. A. and Yi, K.-M. (2006) "Can the standard international business cycle model explain the relation between trade and comovement?" *Journal of International Economics*, Vol. 68, pp. 267–295.

Krishna, V. (1992) "Learning in games with strategic complementarities,"Technical report, Harvard Business School.

Krugman, P. (1996) *The self-organizing economy*: Blackwell Publishers.

Kuramoto, Y. (1975) "Self-entrainment of a population of coupled non-linear oscillators," in *International symposium on mathematical problems in theoretical physics*, pp. 420–422, Springer.

Lane, P. R. and Milesi-Ferretti, G. M. (2007) "The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970–2004," *Journal of International Economics*, Vol. 73, pp. 223–250.

Lecutier, D. (2013) "Stochastic dynamics of game learning," Master thesis, University of Manchester.

Liao, W. and Santacreu, A. M. (2015) "The trade comovement puzzle and the margins of international trade," *Journal of International Economics*, Vol. 96, pp. 266–288.

Long, J. B. and Plosser, C. (1983) "Real business cycles," *Journal of Political Economy*, Vol. 91, pp. 39–69.

Lorenz, H.-W. (1987) "Strange attractors in a multisector business cycle model," *Journal of Economic Behavior & Organization*, Vol. 8, pp. 397–411.

——— (1993) *Nonlinear dynamical economics and chaotic motion*, Vol. 334: Springer.

Lucas, R. E. J. (1977) "Understanding business cycles," in *Carnegie-Rochester conference series on public policy*, Vol. 5, pp. 7–29, North-Holland.

Macy, M. W. and Flache, A. (2002) "Learning dynamics in social dilemmas," *Proceedings of the National Academy of Sciences*, Vol. 99, pp. 7229–7236.

Matsuyama, K. (1999) "Growing through cycles," *Econometrica*, Vol. 67, pp. 335–347.

——— (2007) "Credit traps and credit cycles," *American Economic Review*, Vol. 97, pp. 503–516.

——— (2013) "The Good, the Bad, and the Ugly: An inquiry into the causes and nature of credit cycles," *Theoretical Economics*, Vol. 8, pp. 623–651.

Matsuyama, K., Sushko, I., and Gardini, L. (2014) "Globalization and synchronization of innovation cycles."

May, R. M. (1972) "Will a large complex system be stable?" *Nature*, Vol. 238, pp. 413–414.

——— (1973) "Qualitative stability in model ecosystems," *Ecology*, Vol. 54, pp. 638–641.

Maynard Smith, J. (1982) *Evolution and the Theory of Games*: Cambridge University Press.

McKelvey, R. D. and Palfrey, T. R. (1995) "Quantal response equilibria for normal form games," *Games and Economic Behavior*, Vol. 10, pp. 6–38.

McLennan, A. (2005) "The expected number of Nash equilibria of a normal form game," *Econometrica*, Vol. 73, pp. 141–174.

McLennan, A. and Berg, J. (2005) "Asymptotic expected number of Nash equilibria of two-player normal form games," *Games and Economic Behavior*, Vol. 51, pp. 264–295.

Metzler, L. A. (1941) "The nature and stability of inventory cycles," *Review of Economics and Statistics*, Vol. 23, pp. 113–129.

Milgrom, P. and Roberts, J. (1990) "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica*, pp. 1255–1277.

——— (1991) "Adaptive and sophisticated learning in normal form games," *Games and economic Behavior*, Vol. 3, pp. 82–100.

Miller, R. E. and Blair, P. D. (2009) *Input-output analysis: foundations and extensions*: Cambridge University Press.

Minsky, H. P. (1992) "The financial instability hypothesis," *The Jerome Levy Economics Institute Working Paper*.

Miyazawa, K. (1961) "On the convergence of the learning process in a 2×2 non-zero-sum two-person game,"Technical Report Research Memorandum No. 33, Econometric Research Program, Princeton University.

Monderer, D. and Sela, A. (1996) "A 2×2 Game without the Fictitious Play Property," *Games and Economic Behavior*, Vol. 14, pp. 144–148.

Monderer, D. and Shapley, L. S. (1996a) "Fictitious play property for games with identical interests," *Journal of Economic Theory*, Vol. 68, pp. 258–265.

——— (1996b) "Potential games," *Games and Economic Behavior*, Vol. 14, pp. 124–143.

Mookherjee, D. and Sopher, B. (1994) "Learning behavior in an experimental matching pennies game," *Games and Economic Behavior*, Vol. 7, pp. 62–91.

——— (1997) "Learning and decision costs in experimental constant sum games," *Games and Economic Behavior*, Vol. 19, pp. 97–132.

Morriss, G. P. and Evans, D. J. (2013) *Statistical mechanics of nonequilibrium liquids*: ANU Press.

Mosekilde, E., Larsen, E. R., Sterman, J. D., and Thomsen, J. S. (1992) "Nonlinear mode-interaction in the macroeconomy," *Annals of Operations Research*, Vol. 37, pp. 185–215.

Murphy, K. M., Shleifer, A., and Vishny, R. W. (1989) "Building blocks of market clearing business cycle models," *NBER Macroeconomics Annual*, Vol. 4, pp. 247–287.

Myerson, R. B. (1978) "Refinements of the Nash equilibrium concept," *International Journal of Game Theory*, Vol. 7, pp. 73–80.

——— (2012) "A model of moral-hazard credit cycles," *Journal of Political Economy*, Vol. 120, pp. 847–878.

——— (2013) *Game theory*: Harvard university press.

Nachbar, J. H. (1990) ""Evolutionary" selection dynamics in games: Convergence and limit properties," *International Journal of Game Theory*, Vol. 19, pp. 59–89.

——— (1997) "Prediction, optimization, and learning in repeated games," *Econometrica*, pp. 275–309.

Nagel, R. (1995) "Unraveling in guessing games: An experimental study," *American Economic Review*, Vol. 85, pp. 1313–1326.

Nakao, H. (2016) "Phase reduction approach to synchronization of nonlinear oscillators," *Contemporary Physics*, Vol. 57, pp. 188–214.

Nash, J. F. (1950) "Equilibrium points in n-person games," *Proceedings of the National Academy of Sciences*, Vol. 36, pp. 48–49.

Navas, A., Villacorta-Atienza, J., Leyva, I., Almendral, J., Sendiña-Nadal, I., and Boccaletti, S. (2015) "Effective centrality and explosive synchronization in complex networks," *Physical Review E*, Vol. 92, p. 062820.

Ng, E. C. (2010) "Production fragmentation and business-cycle comovement," *Journal of International Economics*, Vol. 82, pp. 1–14.

Nisan, N., Roughgarden, T., Tardos, E., and Vazirani, V. V. (2007) *Algorithmic game theory*, Vol. 1: Cambridge University Press.

Opper, M. and Diederich, S. (1992) "Phase transition and 1/f noise in a game dynamical model," *Physical Review Letters*, Vol. 69, pp. 1616–1619.

——— (1999) "Replicator dynamics," *Computer Physics Communications*, Vol. 121, pp. 141–144.

Osborne, M. J. and Rubinstein, A. (1994) *A course in game theory*: MIT Press.

Ott, E. (2002) *Chaos in dynamical systems*: Cambridge University Press.

Pangallo, M., Heinrich, T., and Farmer, J. D. (2019) "Best reply structure and equilibrium convergence in generic games," *Science Advances*, Vol. 5, p. eaat1328.

Pangallo, M., Sanders, J. B., Galla, T., and Farmer, J. D. (2017) "A taxonomy of learning dynamics in 2 x 2 games," Available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2913183`.

Papadimitriou, C. and Piliouras, G. (2016) "From Nash equilibria to chain recurrent sets: Solution concepts and topology," in *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science*, pp. 227–235, ACM.

Pecora, L. M. and Carroll, T. L. (1990) "Synchronization in chaotic systems," *Physical Review Letters*, Vol. 64, p. 821.

——— (1998) "Master stability functions for synchronized coupled systems," *Physical Review Letters*, Vol. 80, p. 2109.

Pikovsky, A., Kurths, J., Rosenblum, M., and Kurths, J. (2003) *Synchronization: a universal concept in nonlinear sciences*, Vol. 12: Cambridge University Press.

Powers, I. Y. (1990) "Limiting distributions of the number of pure strategy Nash equilibria in N-person games," *International Journal of Game Theory*, Vol. 19, pp. 277–286.

Rapoport, A. and Guyer, M. (1966) "A taxonomy of 2 x 2 games," *General Systems*, Vol. 11, pp. 203–214.

Rinott, Y. and Scarsini, M. (2000) "On the number of pure strategy Nash equilibria in random games," *Games and Economic Behavior*, Vol. 33, pp. 274–293.

Robinson, J. (1951) "An iterative method of solving a game," *Annals of Mathematics*, pp. 296–301.

Rocchetta, S. and Mina, A. (2019) "Technological coherence and the adaptive resilience of regional economies," *Regional Studies*, pp. 1–14.

Rodrigues, F. A., Peron, T. K. D., Ji, P., and Kurths, J. (2016) "The Kuramoto model in complex networks," *Physics Reports*, Vol. 610, pp. 1–98.

Rosenthal, R. W. (1973) "A class of games possessing pure-strategy Nash equilibria," *International Journal of Game Theory*, Vol. 2, pp. 65–67.

Saari, D. G. (1985) "Iterative price mechanisms," *Econometrica*, pp. 1117–1131.

Sanders, J. B., Farmer, J. D., and Galla, T. (2018) "The prevalence of chaotic dynamics in games with many players," *Scientific reports*, Vol. 8, p. 4902.

Sano, M. and Sawada, Y. (1985) "Measurement of the Lyapunov spectrum from a chaotic time series," *Physical Review Letters*, Vol. 55, p. 1082.

Sato, Y., Akiyama, E., and Farmer, J. D. (2002) "Chaos in learning a simple two-person game," *Proceedings of the National Academy of Sciences*, Vol. 99, pp. 4748–4751.

Sato, Y. and Crutchfield, J. P. (2003) "Coupled replicator equations for the dynamics of learning in multiagent systems," *Physical Review E*, Vol. 67, p. 015206.

Scheinkman, J. A. and LeBaron, B. (1989) "Nonlinear dynamics and stock returns," *The Journal of Business*, Vol. 62, pp. 311–337.

Scheinkman, J. A. (1976) "On optimal steady states of n-sector growth models when utility is discounted," *Journal of Economic Theory*, Vol. 12, pp. 11–30.

Schönhofer, M. (1999) "Chaotic learning equilibria," *Journal of Economic Theory*, Vol. 89, pp. 1–20.

Selover, D. D. and Jensen, R. V. (1999) "Mode-locking and international business cycle transmission," *Journal of Economic Dynamics and Control*, Vol. 23, pp. 591–618.

——— (1991) "Anticipatory learning in two-person games," in R. Selten ed. *Game Equilibrium Models I*, Berlin-Heidelberg: Springer-Verlag, pp. 98–154.

Selten, R. (1975) "Reexamination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, Vol. 4, pp. 25–55.

Shapley, L. S. (1964) "Some topics in two-person games," *Advances in game theory, Annals of Mathematical Studies*, Vol. 52, pp. 1–29.

Shintani, M. and Linton, O. (2003) "Is there chaos in the world economy? A nonparametric test using consistent standard errors," *International Economic Review*, Vol. 44, pp. 331–357.

——— (2004) "Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos," *Journal of Econometrics*, Vol. 120, pp. 1–33.

Shleifer, A. (1986) "Implementation cycles," *Journal of Political Economy*, Vol. 94, pp. 1163–1190.

Silverberg, G. and Lehnert, D. (1993) "Long waves and "evolutionary chaos" in a simple Schumpeterian model of embodied technical change," *Structural Change and Economic Dynamics*, Vol. 4, pp. 9–37.

Skyrms, B. (1992) "Chaos in game dynamics," *Journal of Logic, Language and Information*, Vol. 1, pp. 111–130.

Slutzky, E. (1937) "The summation of random causes as the source of cyclic processes," *Econometrica*, pp. 105–146.

Smith, M. J. (1984) "The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov," *Transportation Science*, Vol. 18, pp. 245–252.

Sonsino, D. (1997) "Learning to learn, pattern recognition, and Nash equilibrium," *Games and Economic Behavior*, Vol. 18, pp. 286–331.

Sparrow, C., Van Strien, S., and Harris, C. (2008) "Fictitious play in $3\times$ 3 games: The transition between periodic and chaotic behaviour," *Games and Economic Behavior*, Vol. 63, pp. 259–291.

Stahl, D. O. (1988) "On the instability of mixed-strategy Nash equilibria," *Journal of Economic Behavior & Organization*, Vol. 9, pp. 59–69.

——— (1996) "Boundedly rational rule learning in a guessing game," *Games and Economic Behavior*, Vol. 16, pp. 303–330.

Stanford, W. (1995) "A note on the probability of k pure Nash equilibria in matrix games," *Games and Economic Behavior*, Vol. 9, pp. 238–246.

——— (1996) "The limit distribution of pure strategy Nash equilibria in symmetric bimatrix games," *Mathematics of Operations Research*, Vol. 21, pp. 726–733.

——— (1999) "On the number of pure strategy Nash equilibria in finite common payoffs games," *Economics Letters*, Vol. 62, pp. 29–34.

Starnini, M., Boguñá, M., and Serrano, M. Á. (2019) "The interconnected wealth of nations: Shock propagation on global trade-investment multiplex networks," *Scientific Reports*, Vol. 9, pp. 1–10.

Sterman, J. D. (1988) "Deterministic chaos in models of human behavior: Methodological issues and experimental results," *System Dynamics Review*, Vol. 4, pp. 148–178.

Strogatz, S. (2004) *Sync: The emerging science of spontaneous order*: Penguin UK.

Suppes, P. and Atkinson, R. C. (1960) *Markov learning models for multiperson interactions*: Stanford University Press.

Tang, F.-F. (2001) "Anticipatory learning in two-person games: some experimental results," *Journal of Economic Behavior & Organization*, Vol. 44, pp. 221–232.

Thorsrud, L. A. (2018) "Words are the new numbers: A newsy coincident index of the business cycle," *Journal of Business & Economic Statistics*, pp. 1–17.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., and De Vries, G. J. (2015) "An illustrated user guide to the world input–output database: the case of global automotive production," *Review of International Economics*, Vol. 23, pp. 575–605.

Topkis, D. M. (1979) "Equilibrium points in nonzero-sum n-person submodular games," *SIAM Journal on Control and Optimization*, Vol. 17, pp. 773–787.

Van Huyck, J. B., Cook, J. P., and Battalio, R. C. (1997) "Adaptive behavior and coordination failure," *Journal of Economic Behavior & Organization*, Vol. 32, pp. 483–503.

Van Huyck, J., Battalio, R., Mathur, S., Van Huyck, P., and Ortmann, A. (1995) "On the origin of convention: Evidence from symmetric bargaining games," *International Journal of Game Theory*, Vol. 24, pp. 187–212.

Vega-Redondo, F. (1996) *Evolution, games, and economic behaviour*: Oxford University Press.

Vives, X. (1990) "Nash equilibrium with strategic complementarities," *Journal of Mathematical Economics*, Vol. 19, pp. 305–321.

Watts, D. J. and Strogatz, S. H. (1998) "Collective dynamics of "small-world" networks," *Nature*, Vol. 393, p. 440.

Woodford, M. (1988) "Imperfect financial intermediation and complex dynamics," in W. A. Barnett, J. Geweke, and K. Shell eds. *Economic complexity: Chaos, sunspots, bubbles and nonlinearity*: Cambridge University Press.

Young, H. P. (1993) "The evolution of conventions," *Econometrica*, Vol. 61, pp. 57–84.

——— (2004) *Strategic learning and its limits*: Oxford University Press.

——— (2009) "Learning by trial and error," *Games and Economic Behavior*, Vol. 65, pp. 626–643.