CERGE-EI
Center for Economics Research and Graduate Education
Charles University Prague

Essays on Economics of Adaptive Learning and Imperfect Monitoring

Batlome Janjgava

Dissertation

Prague, August 2014

Dissertation Committee:

Sergey Slobodyan (CERGE-EI, Chair)

Krešimir Žigic (CERGE-EI)

Petr Zemčik (CERGE-EI)

Referees:

Kenneth Kasa (Simon Fraser University)

Bruce McGough (University of Oregon)

To Juile, Samuel, and my parents.

Acknowledgements

I would like to express my sincere gratitude to my advisor, Professor Sergey Slobodyan, whose advice, encouragement, commitment and enthusiasm helped me to complete my dissertation. I would like to thank Krešimir Žigic, Petr Zemčik, Fabio Michelucci and referees for their insightful comments and suggestions. A special thank goes to the English department, which helped me to improve the readability of the text. I am also grateful to Sharka Cisarova for helping me with the Czech abstract translation. Lastly, I would like to thank CERGE-EI for the knowledge and support that allowed me to pursue the dissertation research.

Abstract

The topic of this dissertation is equilibrium selection in models with incomplete and imperfect information. The dissertation consists of three chapters. In the first two chapters, I focus on firms' decision problems with a structural uncertainty and imperfect monitoring. In the first chapter, co-authored with Sergey Slobodyan, we study a market with two firms competing in quantities. Firms are uncertain about demand parameters and have to learn them using price signals. Although the Cournot output is the Nash equilibrium in the model, we identify conditions when cooperative behavior may arise due to learning and find an endogenous price threshold that triggers such behavior. We show that cooperation is more probable in markets with higher precision of firm-specific shocks.

In the second chapter, I investigate the social efficiency of free entry in homogeneous product markets. In general, free entry is considered desirable for a society from a social welfare point of view and thus, represents traditional wisdom among economic professions. However, many economists have challenged this view and shown that under Cournot oligopoly with fixed setup costs, the free entry equilibrium always delivers excessive entry in homogeneous product markets, known as the excess entry theorem. In this chapter, I reexamine the validity of the excess entry theorem. The theorem advocates restrictive entry policies; nevertheless, I find conditions when free entry is indeed efficient by introducing demand uncertainty into the picture propagating collusive pricing behavior and thus, creating room for additional entry.

In the final chapter, I study the long run outcomes of the belief-based learning process in the infinitely repeated prisoner's dilemma with anonymous random matching and unknown payoff distributions played by a continuum of players. In games with a unique strict Nash equilibrium, such as the prisoner's dilemma, the standard belief-based learning models predict the Nash equilibrium as the only long-run outcome of the learning process. On the other hand, aspiration-based learning models allow dominated strategies to be played in the long run. The opposite predictions of the learning models are often associated with a different level of rationality adopted in the models. However, in this chapter, I show that an important role is, nevertheless, played by informational assumptions. I find that the predictions of the belief-based learning models coincide with the predictions of aspiration-based learning as long as the public signals are perfectly precise and each player puts all weights on those signals. As a result, the only long-run outcome of the learning process is cooperation.

Abstrakt

Tématem této dizertační práce je výběr rovnováhy v modelech s neúplnými a nedokonalými informacemi. Práce se skládá ze tří kapitol, z nichž první dvě zkoumají problémy firem s rozhodováním během strukturální nejistoty a nedokonalého monitorování. První kapitola, jejímž spoluautorem je Sergey Slobodyan, se zabývá trhem, na kterém dvě firmy soutěží nabídkou množství. Firmy si nejsou jisty parametry poptávky a musí je poznávat prostřednictvím cenových signálů. Přestože Cournotův výstup je v našem modelu Nashova rovnováha, identifikujeme nejen podmínky, za kterých může nastat spolupráce z důvodu poznávání, ale rovněž nalézáme endogenní cenový práh, který takovéto chování spouští. Ukazujeme, že spolupráce je pravděpodobnější na trzích s větší přesností šoků specifických pro firmu.

V druhé kapitole se zabývám sociální efektivitou volného vstupu na trhy s homogenními produkty. Volný vstup je zpravidla vnímán jako žádoucí z důvodu společenského blahobytu, díky kterému jej tradičně uznává i ekonomická profese. Mnozí ekonomové však tento názor zpochybňují s ohledem na Cournotovův oligopol s fixními zřizovacími náklady, kdy rovnováha s volným vstupem vždy způsobuje nadměrný vstup na trhy s homogenními produkty, což je známo jako teorém nadměrného vstupu. V této kapitole znovu přezkoumávám platnost teorému nadměrného vstupu. Teorém nadměrného vstupu sice vyžaduje restriktivní vstupní postupy, nicméně za určitých podmínek shledávám volný vstup efektivní, tj. pokud do modelu zavedeme nejistotu ohledně poptávky, a rozšíříme tím koluzivní stanovování cen, které následně vytváří prostor pro další vstup.

V poslední kapitole zkoumám dlouhodobé výsledky učícího procesu založeného na přesvědčení v rámci do nekonečna se opakujícího vězňova dilematu s anonymním náhodným párováním a neznámým rozdělením výnosů hraným kontinuem hráčů. V případě her s jedinou striktní Nashovou rovnováhou, jakou je vězňovo dilema, standardní modely učení založeného na přesvědčení předpovídají, že jediným dlouhodobým výsledkem procesu učení je Nashova rovnováha. Na straně druhé existují modely učení založené na aspiraci, které umožňují hraní dominovaných strategií v dlouhém období. I když jsou protichůdné předpovědi modelů učení často asociovány s různými úrovněmi racionality, v závěrečné kapitole docházím k závěru, že důležitou roli hrají především předpoklady ohledně informací. Předpověď' modelů učení založených na přesvědčení se tak shodují s předpověd'mi modelů učení založených na aspiraci za předpokladu, že každý hráč dostává perfektně přesné veřejné signály a přiřazuje jim veškerou váhu. Jediným dlouhodobým výsledkem procesu učení je pak spolupráce.

Contents

Chapter 1

Duopoly competition, escape dynamics and non-cooperative collusion*

In this paper, we study an imperfect monitoring model of duopoly under similar settings to those in Green and Porter (1984), but in our model firms do not know the demand parameters and learn about them over time through the price signals. We investigate how a deviation from rational expectations affects the decision making process and what kind of behavior is sustainable in equilibrium. We find that the more common information firms analyze to update their beliefs, the more room there is for implicit coordination. This might propagate escapes from the Cournot-Nash equilibrium and the formation of cartels without explicit cooperative motives. In contrast to Green and Porter (1984), our results show that in a model with learning, the breakdown of a cartel happens even without a demand shock. Moreover, in this model an expected price serves as an endogenous price threshold, which triggers a price war. Finally, by investigating the duration of the cooperative and price war phases, we find that in industries with a higher Nash equilibrium output and a lower volatility of firm-specific shocks, it is easier to maintain a cartel and harder to break it down.

Keywords: Beliefs, Escape Dynamics, Implicit Collusion, Self-Confirming Equilibrium, Learning

JEL Classification: D83,D43, L13, L40

*An earlier version of this paper has been published in Janjgava, B. and Solobodyan, S. (2011) Duopoly competition, escape dynamics and non-cooperative collusion, CERGE-EI Working Paper Series, 2011, No. 445. We would like to thank Krešimir Žigic, Levent Çelik and conference participants at the RSA Roma 2011 and at the CEF 2012 for valuable comments and suggestions. All remaining errors are the responsibility of the authors. Batlome Janjgava is grateful for the financial support received from the World Bank Research Fellowship grant.

The difficulty lies not so much in developing new ideas as in escaping from old ones...

- *John Maynard Keynes, The General Theory of Employment Interest and Money, 1935*

1.1. Introduction

■ In this paper, we study an imperfect monitoring model of duopoly where firms do not know the demand parameters and learn about them over time through the price signals. Related literature includes the work of Stigler (1964), Green and Porter (1984), Slade (1989), Williams (2001) and Ellison and Scott (2013). Stigler (1964) argues that a cartel might break down under imperfect information with unobservable firm action as a consequence of secret price cutting, because the rivals fail to distinguish between cheating and a negative demand shock. Green and Porter (1984) extend Stigler's ideas of secret price cutting by introducing imperfect information (only the price is observable, but neither the demand shock nor the rivals' output are observable) into a dynamic model of oligopoly. Green and Porter (1984) show that when structural parameters are known and firms can calculate cooperative and Cournot-Nash equilibrium outcomes, we can have both a collusive and Cournot-Nash equilibrium as a non-cooperative rational expectations equilibrium outcome due to the aggregate demand shock.[1]

However, in practice firms never actually know the demand curve (Balvers and Cosimano, 1990). This implies that firms cannot calculate the Cournot-Nash equilibrium payoff and use it as a benchmark punishment payoff to deter deviations from an agreed upon output. Therefore, when structural parameters are unknown, the only sustainable equilibrium concept is the Nash equilibrium.[2] Nevertheless, we often observe tacit collusion and cooperative behavior in practice.

In this paper, we study the same model as Green and Porter (1984) but with unknown structural parameters. We show that learning the unknown structural parameters itself can create room for implicit coordination and propagation of cooperation without the aggregate demand shock.

The equilibrium concept in our model is the self-confirming equilibrium (SCE) developed

[1] Collusive equilibrium is supported by postulating the existence of a trigger price and a punishment period of a fixed duration, during which firms switch to the Cournot-Nash equilibrium for fixed amount of time whenever the price goes below the trigger price. After the fixed punishment period firms revert to the cartel-like behavior. However, no mechanism that allows coordinating on the trigger price and the punishment duration is provided. Abreu et al. (1986) extends Green and Porter's analysis allowing for more general strategies and shows that trigger strategies are indeed optimal.

[2] Kaneko (1982) shows that in supergames when there is no monitoring and "when a player does not even know what single-period payoffs she has achieved thus far ... the only possible SPEs are strategy profiles that call for single-shot Nash actions in all periods" (Friedman, 2000, p. 65).

and studied by Fudenberg and Levine (1993, 2009) which found widespread application in the models of adaptive learning, cf. Sargent (1999), Williams (2001), Cho et al. (2002), Williams (2003), Kasa (2004), Sargent et al. (2009), and Ellison and Scott (2013). In these models, the adaptive learning mechanism combined with the SCE concept is capable of generating recurrent rapid deviations of beliefs from the SCE followed by a slow return. These recurrent deviations, called 'escape dynamics' after Sargent (1999), were used in the cited papers to explain various economic phenomena such as price volatility, business cycles, disinflation, and currency crises.

The papers most closely related to ours from the adaptive learning literature are by Williams (2001) and Ellison and Scott (2013). Williams (2001) studies a duopoly model of a product market with unknown demand parameters and shows the possibility of escapes from the Nash equilibrium.[3] Ellison and Scott (2013) extend Williams' model to the case of non-renewable resource markets. Both Williams (2001) and Ellison and Scott (2013) use the large deviation theory to characterize the dominant escape path and study the additional volatility of the market price the escape dynamics provide.

Although our model's building blocks are based on Williams (2001) and Ellison and Scott (2013), there are differences in other aspects. We adopt a simple one-dimensional Brownian motion approximation approach, developed in Kolyuzhnov et al. (2014) rather than the large deviation theory to study escape dynamics. Given the simple one-dimensional approach, we could easily identify the belief threshold, which triggers escapes from the Nash equilibrium and provide its analytical solution. In addition, Williams (2001) and Ellison and Scott (2013) do not study the competitive, cooperative and price war phases, whereas we do investigate the dependence of expected duration of these different game phases on the structural parameters of the model. Our analytical results based on the approximation using a one-dimensional Brownian motion are confirmed by the simulations of the original problem.

Uncertainty about demand parameters forces firms to form beliefs about the unknown demand parameters and to update them using Bayes' law upon the arrival of new information. Learning in our model is "perpetual", not ceasing even in the long run, because we assume that the firms believe in the possibility of structural changes and incorporate this belief into their estimation model. Moreover, we mostly interested in the effect of learning on optimal behavior and do not focus on the learning process per se. Firms assume their beliefs to remain unchanged in the future when making decisions and they are not concerned with *experimentation* or *signal-*

[3] A more recent version of Williams (2001), which is available online, no longer provides such analysis.

jamming which are extensively investigated in the IO literature on learning (Grossman et al. (1977), Riordan (1985), Fudenberg and Tirole (1986), Balvers and Cosimano (1990), Mirman et al. (1993), Keller and Rady (1999)). This makes firms in our model anticipated rather than expected profit maximizers in the sense of Kreps (1998).

In our linear model, the shocks are normally distributed, and the use of Bayes' law by the firms to update their belief is equivalent to the Kalman filter (see Balvers and Cosimano (1990), Sargent (1999)). Hence, in our model, firms adopt the Kalman filtering algorithm for belief updating. There are several papers from IO literature that use the Kalman filter for updating beliefs, among them Slade (1989), Slade (1990), and Balvers and Cosimano (1990). From this list, the closest paper to our work is Slade (1989), who presents a way of generating price wars in a price-setting duopoly with unknown demand parameters. In Slade's model unknown parameters are random variables and may change with some small probability, whereas in our model parameters are constant. When demand parameters change, firms start updating their beliefs about them using the Kalman filter, and this propagates price wars. Similarly to Green and Porter (1984), the price war in Slade (1989) arises as a consequence of some exogenous process. The frequency of its occurrence is determined by the demand shocks or shifts in demand parameters, since the history of actions does not alter the distribution of exogenous shocks. We, however, show that even when unknown parameters are constant, price wars could happen endogenously, as a result of belief updating. Our findings are consistent with Ellison (1994), who reports that the Green and Porter model's prediction that price wars happen only due to exogenous shocks in equilibrium is not supported by the data. In our model, price wars occur when the mean actual price falls below the mean forecasted price; thus, the mean forecasted price serves as an endogenous trigger of a price war and firms do not have to commit a priori to some agreed upon price trigger as in Green and Porter (1984).

Another difference that distinguishes our paper from Slade (1989) is as follows. In Slade's model, if the unknown parameters do not change quite often, the learning ceases and the firms can calculate a new stationary Nash equilibrium. In our case learning never ceases; therefore, the firms never know the true values of unknown parameters and thus are never able to calculate the Nash equilibrium of the model. The recurrent escapes from a neighborhood of the Nash equilibrium of our model occur in the direction of the cooperative equilibrium, while in Slade's model, the direction of belief updating is determined by the exogenous shift in demand parameters.

The rest of the paper is structured as follows. We first describe the model and the sustainable equilibrium under Rational Expectations (RE). Then in Section 1.3, we develop the firms' control problem, motivate and define the self-confirming equilibrium (SCE), show that the SCE coincides with the Nash equilibrium (NE), give intuitive insights on possible sources of escapes from the NE, and illustrate them by numerical simulations. In Section 1.4, we formally analyze the mean and escape dynamics of beliefs, derive analytical expressions of the belief trigger and mean exit time from the NE. Next, in Section 1.5, we characterize the nature of belief trajectories and study numerically the duration of the Cooperative Equilibrium (CE) and price war as well as the mean exit time from the NE. Section 2.6 concludes.

1.2. The model setup

■ Consider an oligopolistic industry comprised of two firms producing a single, homogeneous good. It is assumed that firms possess homogeneous production technology with constant marginal cost c. Each period, firm i produces $q_{in} = \hat{q}_{in} + \omega_{in}$ units of production where $\hat{q}_{in}$ is controlled by the firm and ω_{in} a random disturbance. The controllable term $\hat{q}_{in}$ is the expected profit maximizing output while ω_{in} is a firm-specific Gaussian shock, $\omega_{in} \sim N(0, \sigma_2^2)$. The industry faces a linear inverse demand schedule:

$$y_n = a - bQ_n + \omega_n, \tag{1.1}$$

where a and b are positive constants, $Q_n = q_{in} + q_{jn}$ is industry supply, and ω_n is the aggregate demand shock. In order to demonstrate that escapes could happen without an aggregate demand shock, we set its variance to zero.

A firm does not know the intercept (a) and slope (b) of the inverse demand curve. Additionally, it is uncertain regarding rivals' supply. This industry structure may describe the situation in many non-renewable resource extraction industries which are increasingly dominated by global conglomerates, and the information about actual production levels is unreliable in the short run as with, e.g., the OPEC countries' oil production levels.

Because of uncertainty of the structural parameters (a and b), the firms depart from rational expectations and use an adaptive learning algorithm to form expectations instead. The firms attempt to learn the unknown parameters from past observations: they form perceptions about the economic environment, construct beliefs about unknown parameters based on their perceptions and adopt a belief updating mechanism consistent with their belief system. Time varying beliefs affect firms' actions (quantity produced) which in turn affect prices. The firms then use newly available price and quantity data points to update their beliefs.

This interconnection of price signals, beliefs and actions establishes a self-referential process. Because of this self-referential nature under adaptive learning, our equilibrium concept is the self-confirming equilibrium (SCE) developed and studied by Fudenberg and Levine (1993, 2009), which was introduced into the adaptive learning literature by Sargent (1999). Typically, the SCE is defined as a point where the agents' assumption of the orthogonality of the regression error term to regressors is satisfied; therefore, even though the agents' model of their economic environment might be misspecified, at the SCE they do not have an incentive to revise their subjective model which generates "correct" orthogonality prediction. In this paper, the agents will be using a misspecified model in their learning, which makes the SCE an equilibrium concept appropriate for our purposes.

To define the SCE first of all we have to understand how a firm builds its subjective model based on the rational expectations equilibrium of the game. Therefore, in the next section we define the REE of our model and postpone the definition of the SCE till Section 1.3 where we deal with a firm's control problem. In our model, we do not consider strategic manipulations: the firms are assumed to believe that they cannot influence the rivals' decisions even indirectly through the price signals, and thus derive their actions based purely on the currently estimated demand function.

□ **Equilibrium under Rational Expectations.** In this subsection we abstract from learning, equip both firms with rational expectations, and solve for the rational expectation equilibrium (REE) of the model. We consider both the non-cooperative (Nash equilibrium (NE)) and the cooperative (collusive equilibrium (CE)) rational expectations equilibria of our model, and then investigate the stability of each equilibrium in our model settings under RE.

In the Nash equilibrium, each period the firm solves the following profit maximization problem taking the rivals' output as given:

$$\max_{\{\hat{q}_{ik}\}_n^\infty} \mathrm{E}\left[\sum_{k=n}^{\infty} \beta^{k-n}(y_k - c)q_{ik}\right]$$

$$s.t.\ y_k = a - b(q_{ik} + q_{jk}).$$

Solving the above control problem for the symmetric equilibrium provides industry supply and price in the Nash equilibrium:[4]

$$\bar{Q} = \frac{2(a-c)}{3b}, \; \bar{y} = \frac{a+2c}{3}. \tag{1.2}$$

In the collusive equilibrium, the joint expected profit is maximized and the individual firms share it equally since they are assumed to have the same marginal cost:

$$\max_{\{\hat{Q}_k\}_n^\infty} \mathrm{E}\left[\sum_{k=n}^{\infty} \beta^{k-n}(y_k - c)Q_k\right]$$

$$s.t. \; y_k = a - bQ_k,$$

Here, the symmetric collusive equilibrium industry supply and price are given as follows:

$$\tilde{q} = \frac{a-c}{4b}, \; \tilde{y} = \frac{a+c}{2}. \tag{1.3}$$

As is well known, the static joint profit maximization problem is not “stable”. There is always temptation to produce more than the agreed upon output and undercut the price. Friedman (1971) provides the solution in supergames to this instability issue when the actions are observable. He shows that when a certain relation between discount factor and the number of firms holds, the CE can be sustained as an equilibrium by credible punishment threats.

However, we cannot use the oligopoly theory of Friedman (1971) in supergames to solve for the rational expectation equilibrium of the model since here actions are not observable. Moreover, uncertainty about structural parameters makes it impossible for firms to calculate equilibrium payoffs such as monopoly and Cournot-Nash payoffs. This effectively means that firms do not have an equilibrium implied benchmark payoff, such as the Cournot-Nash payoff, to revert to as a punishment to deter deviations from an agreed upon output.Therefore, firms also cannot use the price-trigger strategies of Green and Porter (1984). These considerations make it clear that in our setting only the NE can be sustained as an equilibrium concept.[5]

After considering the REE of the model and reviewing the instabilities involved in creating the cartel, we are ready to introduce learning into the model and consider its implications for equilibrium outcomes.

[4] In the rest of the paper we use $\bar{x}$ and $\tilde{x}$ for the equilibrium values of variable x in the NE and in the CE respectively.

[5] For the same reason, Nash is the equilibrium concept in the models with private information about costs (see for example Chakrabarti (2010) for further discussions).

1.3. Learning the self-confirming equilibrium

□ **The Firm's Control Problem.** Under adaptive learning a firm does not know the true parameters of the inverse demand curve but estimates them from the history of realized prices and own quantities produced. Here, we construct a firm control problem that governs the structure, estimation and consistency of its beliefs.

Every firm understands that the only sustainable equilibrium strategy of the game is to play the Nash strategy and that this is common knowledge. These assumptions lead the firms to postulate the time-invariance of their average beliefs. This establishes them as anticipated profit rather than expected profit maximizers in the sense of Kreps (1998), meaning that if we define a vector $\gamma_{in} = (\gamma^0_{in}, \gamma^1_{in})'$ as a firm i's belief about the intercept and slope of the inverse demand function at the time period n, then the belief evolution equation can be written as

$$\gamma_{in} = \gamma_{in-1} + \eta_{in}, \tag{1.4}$$

where $\eta_{in} \sim N(0, V_i)$. A firm's belief evolution equation (1.4) follows a locally constant random walk process. This type of belief formation is commonly used in economic literature (see Stock and Watson (1996), Sargent (1999), Cho et al. (2002)).

The absence of demand shocks and the time-invariance of beliefs lead firms to form the perception that the rivals' supply is time-invariant on average. This gives firms enough of a basis to build a subjective model that is consistent both with their beliefs and with the REE of the model. To distinguish between the "truth" and the firm's perceptions about the economy, we denote firm i's observed price by y_{in}.[6] Based on the firm's perception of constancy of rivals' supply, the perceived law of motion (PLM) of price is given by the following equation:

$$y_{in} = \gamma^0_{in} + \gamma^1_{in} q_{in} + u_{in}, \tag{1.5}$$

where $\gamma^0_{in} = a - b\hat{q}_{jn}$, $\gamma^1_{in} = -b$, and $u_{in} = -b\omega_{jn}$.

At the beginning of each period, firms determine the controllable part ($\hat{q}_{in}$) of its own supply (q_{in}) by solving the following optimization problem:

$$\max_{\{\hat{q}_{ik}\}_n^\infty} \hat{\mathrm{E}} \left[\sum_{k=n}^{\infty} \beta^{k-n} (y_{ik} - c) q_{ik} | y^{n-1} \right] \tag{1.6}$$
$$s.t. \; y_{ik} = \gamma^0_{ik} + \gamma^1_{ik} q_{ik} + u_{ik},$$

[6] Here, $y_n = y_{in}$ but y_n and y_{in} differ in how the total variation in price is decomposed into endogenous and exogenous components off the equilibrium path.

where β is a common discount factor, y^{n-1} is a history of realized prices up to the period $n-1$, $y^n = \{y_0, y_1, y_2, \dots, y_n\}$ and $\hat{\mathrm{E}}[\cdot]$ denotes the firm's subjective expectation.

The expected profit maximization problem above assumes that a firm treats its beginning-of-period beliefs about the structural parameters as the true ones and solves the control problem (1.6) as if it had rational expectations.[7] This assumption transforms the firm's dynamic optimization problem into a static one. Each period the firm solves the following static problem under the subjective expectations conditional on the history of price signals:

$$\begin{aligned} \hat{\pi}_{in} &= \max_{\hat{q}_{in}} \hat{\mathrm{E}}[(y_{in} - c) q_{in} | y^{n-1}] \\ s.t.\ y_{in} &= \gamma^0_{in} + \gamma^1_{in} q_{in} + u_{in}. \end{aligned} \tag{1.7}$$

The solution yields the firm's supply function dependent on its current beliefs:

$$\hat{q}_{in} = \max\left\{ \frac{\hat{\mathrm{E}}[\gamma^0_{in} | y^{n-1}] - c}{-2\,\hat{\mathrm{E}}[\gamma^1_{in} | y^{n-1}]}, 0 \right\}. \tag{1.8}$$

A firm understands that its own action depends on its beliefs about the inverse demand function parameters and needs to adopt an estimation model to estimate them. Based on its perceptions and control problem, the firm builds the following estimation model:

$$\begin{aligned} y_{in} &= x'_{in} \gamma_{in} + u_{in}, \\ \gamma_{in} &= \gamma_{in-1} + \eta_{in} \end{aligned} \tag{1.9}$$

where $x_{in} = (1, q_{in})'$, $u_{in} \sim N(0, \sigma^2)$ and $\sigma^2 = b^2 \sigma^2_2$.

Each period, after the production decision is made, firms observe the price y_n and use Bayes' law to update their beliefs $\hat{\gamma}_{in+1} = \hat{\mathrm{E}}[\gamma_{in+1} | y^n]$ about the unknown parameters using their estimation model (1.9). The optimal Bayesian updating framework, based on the estimation model (1.9), is provided by the Kalman filtering technique and is described by the following recursive algorithm:

$$\begin{aligned} \hat{\gamma}_{in+1} &= \hat{\gamma}_{in} + \frac{\hat{P}_{in}}{1 + x'_{in} \hat{P}_{in} x_{in}} x_{in} (y_{in} - x'_{in} \hat{\gamma}_{in}), \\ \hat{P}_{in+1} &= \hat{P}_{in} - \frac{\hat{P}_{in} x_{in} x'_{in} \hat{P}_{in}}{1 + x'_{in} \hat{P}_{in} x_n} + \sigma^{-2} V_i, \end{aligned} \tag{1.10}$$

where $\hat{P}_{in} = \sigma^{-2} \mathrm{cov}[\gamma_{in} - \hat{\gamma}_{in}]$.[8]

[7] In other words, under $\hat{\mathrm{E}}[\cdot]$ a firm fixes its future belief vectors at the current estimated value.

[8] Note that such behavior is optimal only if the firm does not take into account the effect of its own learning on its

□ **Self-confirming Equilibrium.** The firm's policy rule (1.8) under the PLM defines the actual law of motion (ALM) of the price given the true relationship (2.1) between price, y_n, and industry supply as follows:

$$y_n = a - b[q_{in}(\hat{\gamma}_{in}) + q_{jn}(\hat{\gamma}_{jn})]. \tag{1.11}$$

Every firm treats the subjective model (2.2) it has in mind as if it were the true one. Therefore, they take the errors u_{in} to be orthogonal to the regressors x_{in}. This orthogonality is generally not observed in the data. However, there exists an equilibrium defined as a time-invariant vector of parameters $\bar{\gamma}_i$ for which the assumed orthogonality condition is satisfied, as the true mathematical expectation with respect to ALM of price (2.8) is indeed zero:

$$\mathrm{E}[x_{in}(y_n - x'_{in}\bar{\gamma}_i)] = 0. \tag{1.12}$$

The Self-Confirming Equilibrium is defined as the unique symmetric vector of beliefs in which both firms' orthogonality assumptions are confirmed by observations.

Taking expectations in (1.12) produces the following expression:

$$\begin{bmatrix} \gamma_i^0 \\ \gamma_i^1 \end{bmatrix} = \begin{bmatrix} a - bE[q_{jn}] + \rho b \mathrm{E}[q_{in}] \\ -b - \rho b \end{bmatrix}, \tag{1.13}$$

where $\rho = \mathrm{cov}[q_{jn}, q_{in}] / \mathrm{var}[q_{in}]$.

In the SCE, the beliefs are fixed and thus there is no correlation between the firms' actions, therefore $\rho = 0$. A unique self-confirming equilibrium is then given by

$$\bar{\gamma}_i = \bar{\gamma}_j = (\bar{\gamma}_i^0, \bar{\gamma}_i^1)' = (\bar{\gamma}_j^0, \bar{\gamma}_j^1)' = \left(\frac{2a+c}{3}, -b\right)'. \tag{1.14}$$

From (1.8), at SCE both firms produce

$$\bar{q} \equiv \hat{q}_{in}(\bar{\gamma}) = \hat{q}_{jn}(\bar{\gamma}) = \frac{a-c}{3b}. \tag{1.15}$$

The SCE thus coincides with the Nash equilibrium (1.2). The reason for this is the fact that both equilibria are derived assuming that a firm considers rivals' actions to be independent of its own actions and constant over time. Therefore, deviation from the NE is equivalent to deviation from the SCE. These two concepts will be used interchangeably in what follows in the paper.

future beliefs and prices. In terms of Wieland (2000), the firms are engaged in "passive" learning.

The expression for ρ in (2.9) represents the linear estimate of the conjectural variation that measures the degree of rivals' reaction to the firm's action. The belief that conjectural variation is zero eliminates any incentives to deviate from the SCE.[9] However, if conjectural variations were non-zero, the beliefs would start to deviate from the SCE. A particular form of fast and sudden deviations from SCE is called "escapes" in economic literature, after Sargent (1999).

Since the beliefs are estimated, they are a vector random variable. As both firms use the same price history and have the same PLM (2.2) to update beliefs, there are common factors in the forecast errors made by firms. These common factors allow the firms' beliefs to become correlated. The correlation of beliefs implies correlation of actions and might become a source for escaping from the SCE. Since firms are homogeneous and use the same price signal to update their beliefs we expect ρ to be non-negative during an escape.

As mentioned above, the escapes are fast, which can be explained by the presence of a following positive feedback loop. If the firms' beliefs become correlated sufficiently strongly (ρ is sufficiently positive), the correlation between firms' actions increases so much that the beliefs are pushed even further from the SCE (ρ becoming even larger). Intuitively, we can think of the limit point of the escape as the point where firms' actions get perfectly correlated; $\rho = 1$. As is known from IO literature, ρ equal to one corresponds to the cooperative equilibrium. Beliefs that satisfy the orthogonality condition (1.12) with $\rho = 1$ have the form

$$\tilde{\gamma}_i = \tilde{\gamma}_j = (\tilde{\gamma}_i^0, \tilde{\gamma}_i^1)' = (\tilde{\gamma}_j^0, \tilde{\gamma}_j^1)' = (a, -2b)'. \tag{1.16}$$

Given this belief vector each firm's produced output equals the CE output:

$$\tilde{q} \equiv \hat{q}_{in}(\tilde{\gamma}) = \hat{q}_{jn}(\tilde{\gamma}) = \frac{a-c}{4b}.$$

Thus, we can think of escapes as movements from the Nash equilibrium towards the cooperative equilibrium.

1.4. Escape dynamics

■ In this section, we describe formally the most probable escape path and show that, indeed, escapes happen towards the direction of cooperative equilibrium. We derive the price thresholds that trigger the escapes from the Nash equilibrium towards the collusive equilibrium and vice

[9] This point is proved in a more rigorous way in Section 1.4, where we derive E-stability conditions of the SCE.

versa. Then, we characterize the mean escape time analytically and investigate the factors affecting the frequency of escapes. To study escape dynamics we draw on the approach developed in Kolyuzhnov et al. (2014).

In Section 1.3, the Kalman filter updating scheme was provided as the belief updating mechanism. Another updating mechanism often used in the literature in models with structural changes is constant gain recursive least squares learning (CG RLS), which is equivalent to weighted least square (WLS) with the weights geometrically falling with the age of data points. In order to facilitate the comparison with Ellison and Scott (2013), who use CG RLS, we follow Sargent and Williams (2005) in designing a Kalman filter algorithm that asymptotically has the same expected dynamics as the CG RLS. This equivalence is achieved if the belief updating scheme (1.10) is approximated as follows:

$$\begin{aligned} \hat{\gamma}_{in+1} &= \hat{\gamma}_{in} + \epsilon P_{in} x_{in} u_{in}, \\ \hat{\gamma}_{jn+1} &= \hat{\gamma}_{jn} + \epsilon P_{jn} x_{jn} u_{jn}, \\ P_{jn+1} &= P_{jn} + \epsilon\left(\sigma^{-2}\hat{V}_i - P_{in} M_i(\gamma_{in}) P_{in}\right), \\ P_{jn+1} &= P_{jn} + \epsilon\left(\sigma^{-2}\hat{V}_j - P_{jn} M_j(\gamma_{jn}) P_{jn}\right), \end{aligned} \tag{1.17}$$

and matrix $\widehat{V}_i$ is selected so that $\widehat{V}_i = \sigma^2 M_i(\overline{\gamma}_i)^{-1}$. Here $u_{in} = y_n - x'_{in}\hat{\gamma}_{in}$ is a forecast error, ϵ the gain in a constant gain algorithm, $M_i(\gamma_{in}) = \mathrm{E}[x_{in}x'_{in}]$, and $P_{in} = \epsilon^{-1}\hat{P}_{in}$. The expressions for j are analogous. Thus the approximate Kalman filter defined is asymptotically equivalent to the constant gain recursive least squares algorithm with gain ϵ. Evans et al. (2010, Proposition 1, 2) show that the constant gain adaptive learning asymptotically approximates the optimal Bayesian estimation procedure if the beliefs are given by(1.4). Even if the firm believes the parameters γ_i to be constant over time, a form of the constant gain learning rule could be shown to be the (maximally) robust optimal predictor.

Simulating the system (2.7), we observe in Figure 1.1 that if the beliefs escape the Self-Confirming Equilibrium, they escape towards the cooperative equilibrium and then return back to the SCE. Figure 1.2 shows that escape from the SCE $\bar{\gamma} = (1.333, -0.100)'$ (see (1.14)) happens towards the belief vector $\tilde{\gamma} = (2.000, -0.200)'$ (see (1.16)) that corresponds to full collusion. This confirms our intuition in Section 1.3 about the limit point of escape in the belief space. Moreover, the escape dynamics towards the cooperative equilibrium lies along the line that connects the cooperative and the Nash equilibrium beliefs.

The observed escapes from the SCE are not a result of the instability of the belief updating

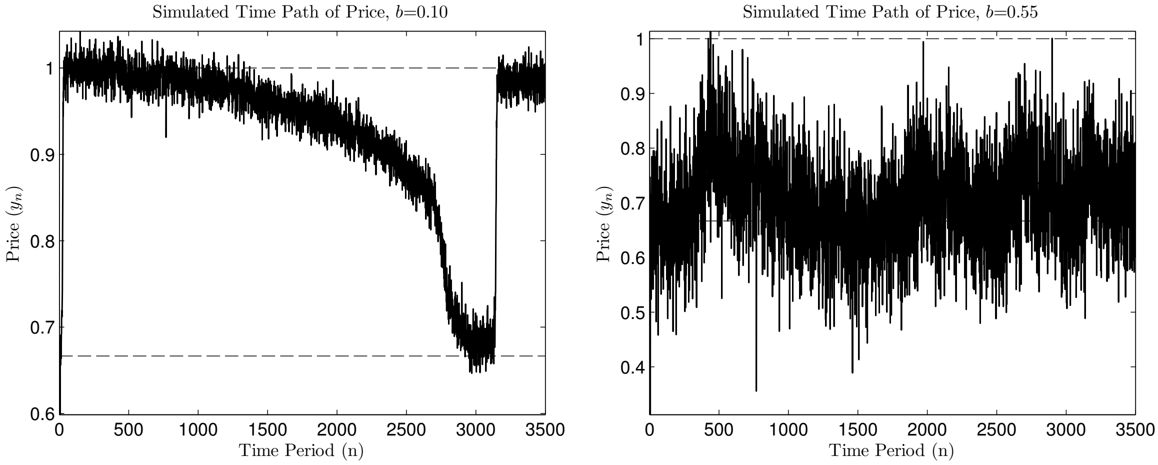


FIGURE 1.1: Simulated time path of price for different specifications of the slope parameter, $a = 2$, $\sigma_2^2 = 0.01$, $\epsilon = 0.005$.

process described by (2.7). To elaborate on this let us consider an average dynamics of (2.7) which is constructed as follows. Fix the beliefs $\gamma_{i,j}$ and $P_{i,j}$. Generate a process of prices y_n that would be obtained for these beliefs from equation (2.8). Then, take the expectation of the right-hand side of (2.7) with respect to the probability distribution induced by the derived stochastic process for y_n. Repeat this process for all possible beliefs and form the following system of ordinary differential equations (ODE):

$$
\begin{aligned}
\dot{\gamma}_i &= P_i \bar{g}_i(\gamma_i, \gamma_j) \\
\dot{\gamma}_j &= P_j \bar{g}_j(\gamma_i, \gamma_j) \\
\dot{P}_i &= \sigma^{-2} \hat{V}_i - P_i M_i(\gamma_i) P_i, \\
\dot{P}_j &= \sigma^{-2} \hat{V}_j - P_i M_j(\gamma_j) P_j,
\end{aligned}
\tag{1.18}
$$

where $\bar{g}_i(\gamma_i, \gamma_j) = \mathrm{E}[x_{in}(y_n - x'_{in}\gamma_i)]$.

As discussed in Evans and Honkapohja (2001), solution paths of the resulting system of ODE approximate the behavior of the real-time learning dynamics (which takes place in discrete time) asymptotically as the gain parameter goes to zero. For more details on deriving the approximating ODE for the Kalman filter that is asymptotically equivalent to the constant gain RLS, consult Sargent and Williams (2005). The system (1.18) is referred to as the mean dynamics ODE, and its solution paths are called mean dynamics trajectories or simply mean dynamics. By construction, mean dynamics approximates the average behavior of the firm's belief when the updating process is given by (2.7).

A unique steady state of (1.18) is given by $\gamma_i^* = \gamma_j^* = \bar{\gamma}$ and $P_i^* = P_j^* = \bar{P} = M(\bar{\gamma})^{-1}$, which

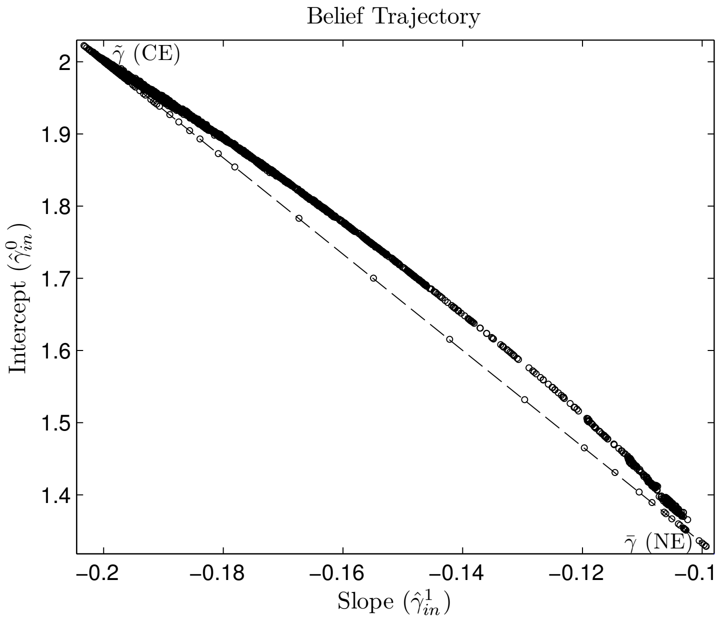


FIGURE 1.2: Simulated belief trajectory (circle) and line connecting CE and NE beliefs (dashed), $a = 2$, $b = 0.1$, $\sigma_2^2 = 0.01$, $\epsilon = 0.005$.

coincides with the SCE. In Appendix A, it is shown that the steady state of the system (1.18) is stable for any parameter values; we say that the SCE beliefs are E-stable. Evans and Honkapohja (2001) show that if a stationary point of the mean dynamics ODE is stable, then the corresponding real-time dynamics with a decreasing gain converge to the stationary point with a probability that could be made as close to one as desired. For initial beliefs within a sufficiently small neighborhood of the SCE, the firms on average will have no incentives to deviate from the Nash equilibrium. Instead, if the firms are using constant gain learning as in this paper, they will not be able to learn the beliefs $\gamma^*_{i,j} = \bar{\gamma}$ and $P^*_{i,j} = M(\bar{\gamma})^{-1}$ exactly, but their beliefs will converge to an invariant distribution centered on $(\bar{\gamma}, \bar{P})$. Given E-stability of the SCE, we can say that the observed escapes are not due to its instability.

Simulations in Figure 1.1 clearly show that for some parameter values the beliefs may escape the SCE towards cooperative equilibrium. It is likely, therefore, that a threshold exists such that an escape from the SCE is triggered when the beliefs reach this threshold. We turn next to finding this belief threshold and studying the likelihood of escape from the SCE as a function of the structural parameters of the model.

□ **Analysis of Mean Dynamics and Trigger of Cooperation.** Since firms are homogeneous and their mean dynamics equations are the same, in order to find the belief threshold, it suffices to restrict our analysis to a symmetric case where $\gamma_j = \gamma_i$. We then analyze the following mean dynamics equations:

$$\begin{aligned}\dot{\gamma} &= P\bar{g}(\gamma),\\ \dot{P} &= \sigma^{-2}\hat{V} - PM(\gamma)P,\end{aligned} \tag{1.19}$$

where the belief vector γ is two-dimensional.

Following Kolyuzhnov et al. (2014), the escapes in the system (2.13) occur predominantly along the direction of the dominant eigenvector $\bar{v}$ corresponding to the largest eigenvalue of $\bar{P}$.[10] We restrict our attention to the line that starts at the SCE and extends in the direction $\bar{v}$. In parametric form, this line is given by the following expression:

$$\Gamma = \{\gamma \mid \gamma = \bar{\gamma} + \delta\bar{v}, \delta \in \mathbb{R}\}. \tag{1.20}$$

Because the dominant eigenvalue of $\bar{P}$ ($\bar{\lambda}_1 = 4545.40$) is significantly larger than the second one ($\bar{\lambda}_2 = 0.02$), solution paths of (2.13) along the escape route are likely to lie very close to Γ. Therefore, Γ represents the dominant escape path.

The escape vector $\bar{v}$ can be written as $\left((1-1/\bar{\lambda}_1)^{-1}\overline{q}, -1\right)^T$, where $\bar{\lambda}_1$ is the largest eigenvalue of $\bar{P}$. The direction $\bar{v}$ is closely aligned with the vector $\tilde{\gamma} - \bar{\gamma}$, whenever $\bar{\lambda}_1$ tends to infinity. The latter can happen whenever σ_2^2 tends to zero. Therefore, the likelihood of observing escapes from the Nash equilibrium (beliefs given by $\bar{\gamma}$) towards exactly the cooperative equilibrium (beliefs $\tilde{\gamma}$) is higher the lower is the volatility of firm-specific shocks.

The belief updating procedure minimizes volatility of a forecast error u_{in}, which suggests that the belief updating is sensitively dependent on the behavior of the mean forecast error. Therefore, to find the belief trigger, we should understand how beliefs γ and the mean forecast error $u = \mathrm{E}[u_{in}]$ interact with each other. Since the right-hand side of the mean dynamics equations (1.18) depends on $\bar{g}_i(\gamma_i, \gamma_j) = E[x_{in}u_{in}]$, the mean forecast error can be used to judge whether the belief updating process is accelerating or stabilizing. When the mean forecast error is growing fast the belief updating accelerates, while we should observe the opposite when the growth of the mean forecast error slows down. Since we observe escapes from the SEC, even

[10] If the constant gain parameter ϵ is not too small, the behavior of equation (2.7) is almost one-dimensional because $\bar{\lambda}_1$, the dominant eigenvalue of $\overline{P}$, is significantly larger than the second one. As a result, increments in the equation (2.7), given by

$$P_n g(\gamma_n, \xi_n) = \sum a_i \lambda_i v_i \approx a_1 \lambda_1 v_1$$

are concentrated in a narrow conus formed around v_1, the dominant eigenvector of $\bar{P}$ (the eigenvalues and eigenvectors of P_n are essentially the same as those of $\bar{P}$ near the SCE). For similar ideas one can also consult the method of principal component analysis.

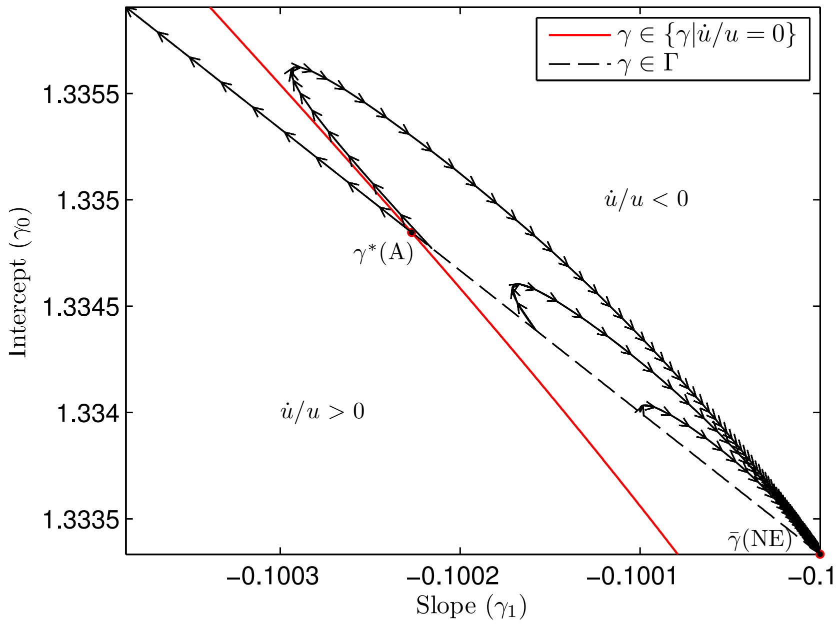


FIGURE 1.3: Phase diagram of beliefs, $a = 2$, $b = 0.1$, $\sigma_2^2 = 0.01$. Two types of mean belief trajectories are depicted: one with initial beliefs inside the attraction region of SCE along the dominant escape path and another for initial beliefs outside the attraction region of SCE along the dominant escape path. Mean belief trajectories are shown by arrows.

though the SCE is a stable steady-state of the system (1.18), there should exist a neighborhood of the SCE such that the mean forecast error is declining over time ($\dot{u}/u < 0$) within it and growing ($\dot{u}/u > 0$) outside of it.

The above logic is illustrated in Figure 2.3 where mean belief trajectories are plotted with different initial beliefs $\gamma \in \Gamma$ together with the curves $\dot{u} = 0$ and $u = 0$ near the SCE. We see that on the segment ANE of line Γ, the mean forecast error is declining; $\dot{u}/u < 0$. If the initial beliefs are located in this segment, the average forecast error is expected to decline, meaning that beliefs are expected to move towards the SCE. However, for an initial belief $\gamma \in \Gamma$, outside of the ANE segment, the beliefs diverge from the SCE since $\dot{u}/u > 0$, and they are expected to move along Γ towards the CE. We see that the neighborhood of $\bar{\gamma}$, where $\dot{u}/u < 0$, behaves like the region of attraction of the SCE along the escape direction; outside of this region the beliefs escape the Self-Confirming Equilibrium.

When firms escape the Nash equilibrium and start cooperating, the industry supply shrinks. Therefore, the belief threshold γ^* that triggers cooperation, and at the same time escape from the SCE, should satisfy the following equation:

$$\begin{cases} \dot{u} = \nabla u(\gamma)\, \dot{\gamma} > 0, \\ \dot{Q} = \nabla Q(\gamma)\, \dot{\gamma} < 0. \end{cases}$$

The $\dot{Q} < 0$ restriction is redundant because the escape of beliefs from the SCE implies a cooperative behavior: for $\gamma \in \Gamma$, $\dot{u} > 0$ implies $\dot{Q} < 0$. This implies that we can find the belief threshold $\gamma^* \in \Gamma$ by solving an algebraic equation $\dot{u} = 0$ for δ^*:

$$\dot{u} = \nabla u(\bar{\gamma} + \delta^* \bar{v})\overline{P}\bar{g}(\bar{\gamma} + \delta^* \bar{v}) = 0.$$

This yields the following approximate expression for δ^*:

$$\delta^* \approx 6\frac{\sigma_2^2}{\bar{q}^2}b. \tag{1.21}$$

The analysis above and in the next section relies on the one-dimensional nature of escaping beliefs to a significant degree. In the case of a strictly one-dimensional dynamics, we could state the following proposition regarding the direction and the size of deviation which triggers escape from the neighborhood of the SCE.

Proposition 1. *The escape from the SCE should occur in the direction given by the vector $\bar{v}$. The escape happens if the beliefs deviate from the SCE by more than $\delta^* \|\bar{v}\|$.*

Because the true dynamics is not one-dimensional, both results in Proposition 1 are correct only asymptotically, as the constant gain ϵ converges to zero. How good is the assumption of one-dimensionality? In Figure 1.4 we plot histograms of directions of escape from the ball $E = \{\gamma \mid \|\gamma - \bar{\gamma}\| \leq \delta^* \|\bar{v}\|\}$ for different values of the constant gain ϵ.[11] For every escape point γ^{esc}, we write $\gamma^{esc} - \bar{\gamma}$ as a vector proportional to $(\kappa, 1)$. If the escapes happen exactly along the direction $\bar{v}$, κ equals -6.668. As is clear from the graph, simulated escapes are distributed symmetrically around the theoretical direction; moreover, the distribution of escape directions becomes concentrated around the mean as the value of the constant gain ϵ declines (Var(κ) is increasing with ϵ). Therefore, considering the escape dynamics as essentially one-dimensional is a valid approximation and could be used to derive the other values of interest.

□ **Time Until Cooperation and Comparative Statics.** In this section we use the expression for δ^* to characterize the expected time until an escape to cooperation happens and investigate its dependence on the structural parameters.

Definition 1. *The mean exit time from the NE is defined as a mean value of N^ϵ where $N^\epsilon = \inf\{n \mid \|\gamma_n - \bar{\gamma}\| > \|\gamma^* - \bar{\gamma}\|\}$.*

[11] The boundary of this ball, a sphere $\{\gamma \mid \|\gamma - \bar{\gamma}\| = \delta^* \|\bar{v}\|\}$, intersects the line Γ at a distance $\delta^* \|\bar{v}\|$ from the SCE.

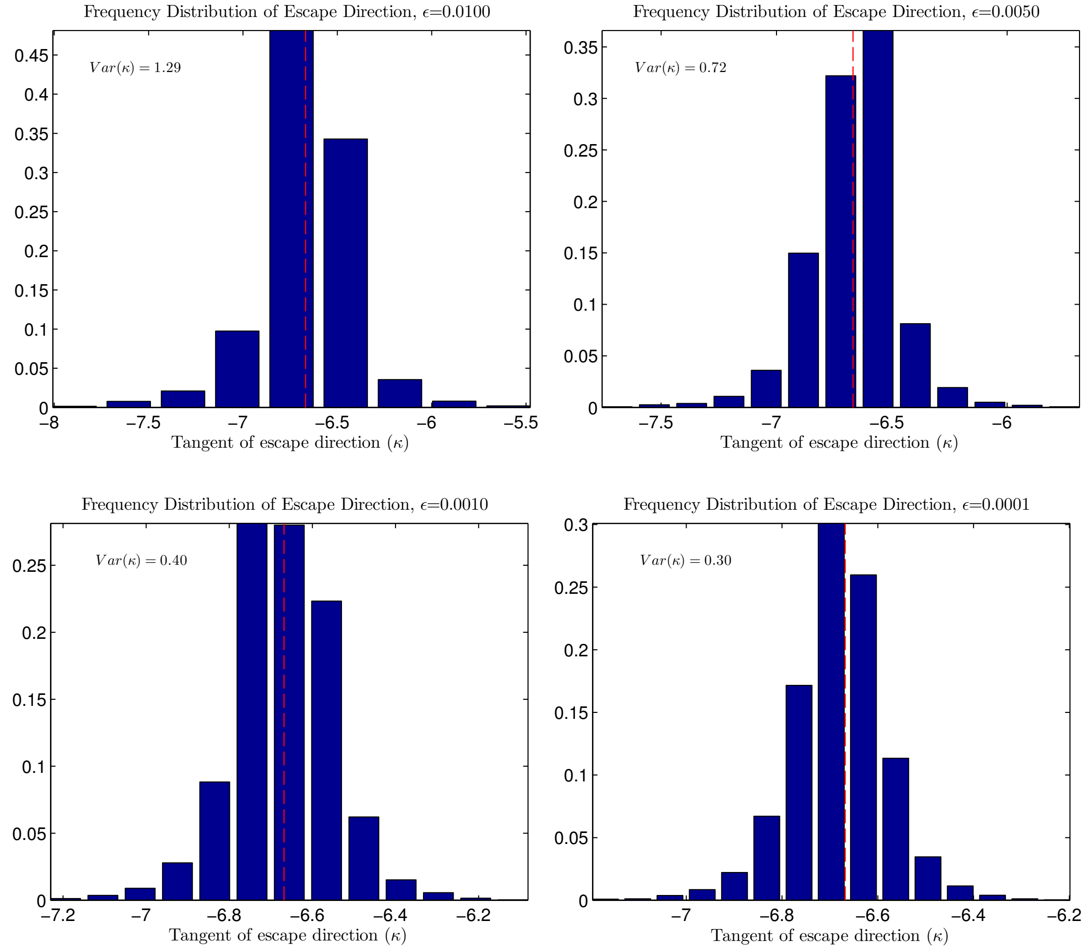


FIGURE 1.4: Frequency distribution of directions of escape from the ball $E = \{\gamma \mid \|\gamma - \bar{\gamma}\| \leq \|\gamma^* - \bar{\gamma}\|\}$ and the direction of $\bar{v}$ (dashed), $a = 2$, $b = 0.1$, $\sigma_2^2 = 0.01$.

The mean exit time is the expected time until the beliefs leave the ball of radius $\|\gamma^* - \bar{\gamma}\|$ around the SCE for the first time. This ball includes the region of attraction of the SCE. After this time, a very fast transition to cooperation follows.

In addition to the mean dynamics approximation (1.18), it is possible to approximate the real-time dynamics under adaptive learning by the following one-dimensional continuous time diffusion process as described in Appendix B:[12]

$$d\hat{\varphi}_t = -\hat{\varphi}_t dt + \sqrt{\epsilon \lambda_{\bar{\Sigma}}} dz_t, \tag{1.22}$$

where $\hat{\varphi}_t = \hat{v}^T(\gamma_t - \bar{\gamma})/\|\hat{v}\|$, $\gamma_t = \left(\gamma_{in}^T, \gamma_{jn}^T\right)^T$, $\lambda_{\bar{\Sigma}} = 2b^2\sigma_2^2\bar{\lambda}_1$, z_t is a one-dimensional Wiener process.

This approximation allows us to derive an analytical expression for the expected number of real-time periods till escape, which is given in the following proposition.

[12]The diffusion process (1.22) is a special case of the well-known Ornstein-Uhlenbeck process.

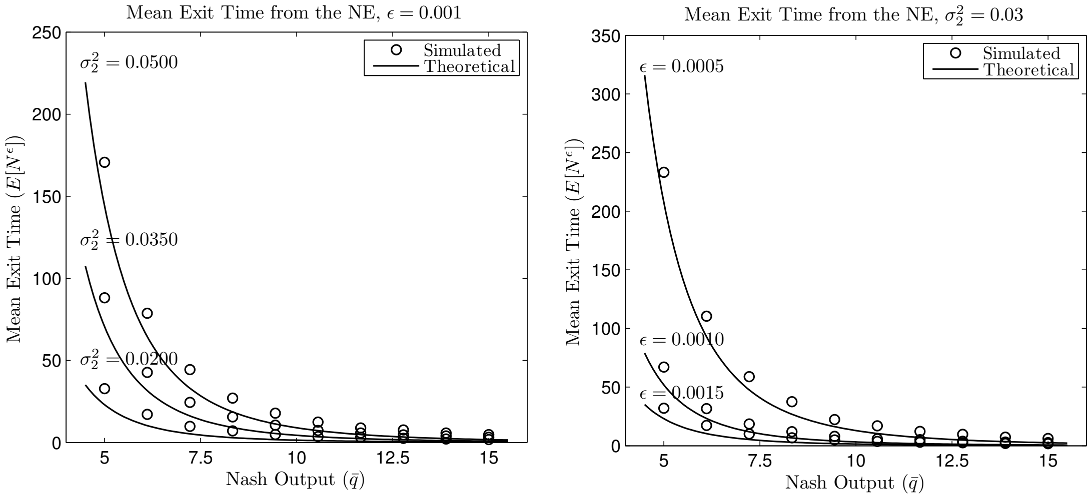


FIGURE 1.5: Comparison of the analytically derived and simulated mean exit times from the NE for different specifications of the model parameters.

PROPOSITION 2. *The expected number of periods until an escape towards collusion happens is inversely proportional to* $\left(\bar{q}/\sqrt{\sigma_2^2}\right)^4$*, the fourth power of the standardized Nash output, and inversely proportional to* ϵ^2*, the square of the gain parameter:*

$$E\left[N^\epsilon\right] \sim \left(\frac{\bar{q}}{\sqrt{\sigma_2^2/\epsilon}}\right)^{-4}. \tag{1.23}$$

We find that the likelihood of cooperation (escape to cooperation is faster) is higher in industries with a higher Nash output $\bar{q}$ and a lower volatility σ_2^2 of firm-specific shocks. This finding is rather intuitive. In industries with a higher Nash output $\bar{q}$, the difference between the competitive and cooperative profits is greater.[13] The latter makes collusion more attractive, leading firms to quit the competitive Nash equilibrium more easily: the expected number of periods the competition could be supported becomes smaller. Since the firm-specific shocks are independent, it is obvious that the higher their volatility, the less correlated firms' actions become. With a lower firm-specific shock variance, the share of uncertainty in the data generating process directly attributed to idiosyncratic shocks decreases, therefore increasing the role of uncertainty related to belief updating which is common among firms. Increasing the share of the common factor raises the probability of firms' beliefs becoming correlated and thus the likelihood of escape from the NE; expected escape time drops.

[13] At NE, firm's average output is $\bar{q}$ and the price $\frac{a}{3}$, giving an average profit of $\frac{a^2}{9b}$. At the CE, the average profit is $\frac{a^2}{8b}$. Therefore, the profit difference is proportional to $\frac{a^2}{9b} = \bar{q}^2 b$. (Here, for simplification, the marginal cost c is set to zero.)

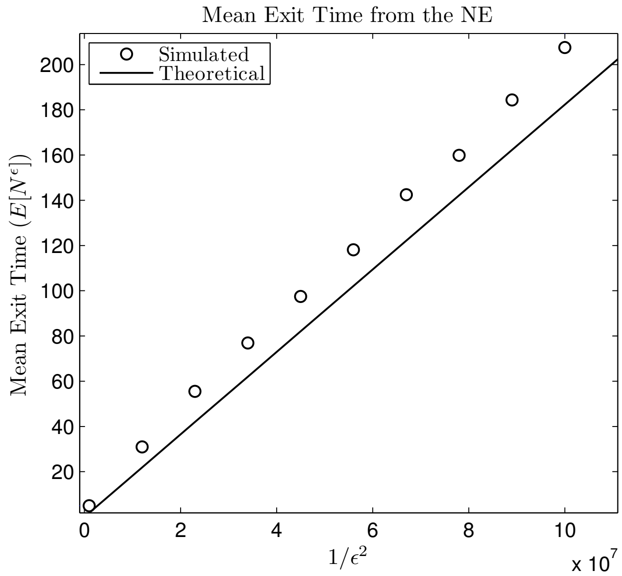


FIGURE 1.6: Linearity of simulated mean exit time from the NE in $1/\epsilon^2$, $a = 2$, $b = 0.1$, $\sigma_2^2 = 0.01$. It shows that the slopes of simulated and analytically derived mean exit times in the space of $1/\epsilon^2$ are quite close to each other.

Finally, the gain parameter ϵ measures the speed with which the firms incorporate the new information into their beliefs; lower ϵ means that the firms update their beliefs more slowly, and it takes a longer stretch of accidental correlation to lead to the firms' reaction.[14] Therefore, lower ϵ makes the competitive phase in the industry last longer.

The above discussion considered an approximated value of the expected escape time that was derived using a one-dimensional approximation. Below we compare this analytical approximation with the simulation evidence which demonstrates that the agreement is remarkably good. In Figure 1.5 we present the average escape time obtained by simulating the original discrete-time learning dynamics (2.7). One could see that the mean exit time from the Nash equilibrium is increasing in the firm-specific shock volatility σ_2^2 and decreasing in Nash output $\overline{q}$ as well as in the constant gain parameter ϵ, exactly as predicted by the analytical expression (1.23). Additionally, the analytically derived and simulated mean exit times are remarkably close.

To illustrate even better the validity of our analytical approximation (1.23), in Figure 1.6 we plot the expected escape time, expressed as a number of periods, against $1/\epsilon^2$. Proposition 2 states that this dependence is given by a straight line with a certain slope. The graph indi-

[14]As another possibility, think about the firms which could alternatively use constant gain updating equations (2.7) with an infinite amount of data or run an OLS regression with T data points. In the OLS regression, the mean age of a data point is $T/2$. A constant gain learning algorithm is equivalent to a weighted Least Squares, where the weight of a point τ periods old is proportional to $(1-\epsilon)^\tau$, giving $\frac{1-\epsilon}{\epsilon}$ as the mean age of a point when an infinite amount of data is used. Thus, $\epsilon \approx 2/T$ produces a similar average age of the data for the OLS and the constant gain RLS. Therefore, lower ϵ is equivalent to using more data in forming the beliefs by OLS or forgetting this data at a slower rate in constant gain RLS.

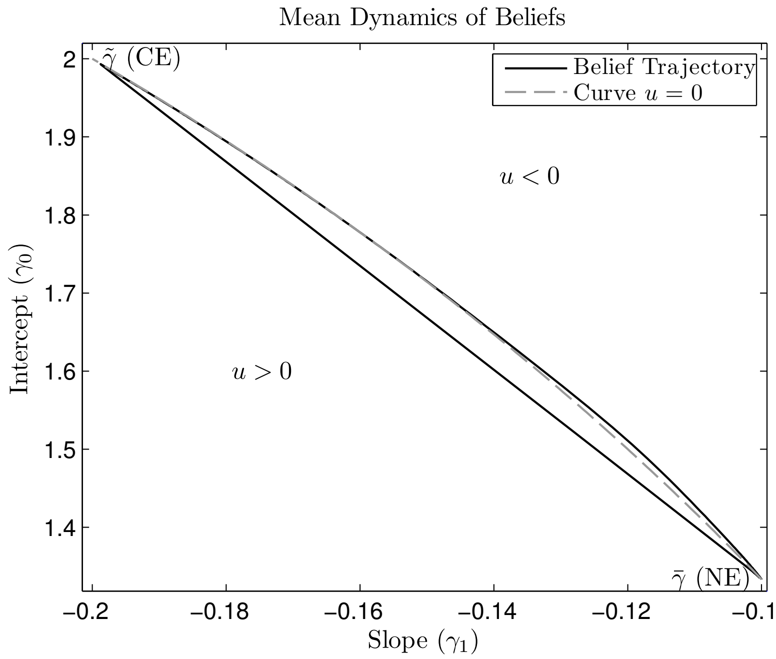


FIGURE 1.7: Comparison of the mean belief trajectory and zero mean forecast error curve ($u = 0$), $a = 2$, $b = 0.1$, $\sigma_2^2 = 0.01$, $\epsilon = 0.005$.

cates, first, that average simulated escape times are indeed scaled as $1/\epsilon^2$. Second, the observed slope is remarkably close to the theoretically derived one. The figure unambiguously demonstrates that the escape dynamics could indeed be derived from the behavior of one-dimensional approximation with great precision.

1.5. Characterizing the belief trajectories

■ In this section we first characterize belief trajectories and define cooperative and price war phases. Next, we do numerical analysis of durations of these phases and their functional dependence on the model parameters.

□ **From Competition to Cooperation and Back: Forecast Errors.** We can notice from Figure 1.1 and Figure 1.2 that the beliefs are characterized by fast and slow dynamics. When the beliefs escape the Nash equilibrium, the firms quickly learn how to cooperate and rush towards the cooperative equilibrium: it can take less than ten periods from the time beliefs leave the SCE region of attraction to their arrival in the neighborhood of the CE. After the cooperative equilibrium is reached, the dynamics slow down dramatically. Below we give the reason for these fast and slow dynamics and associate them to the cooperative and price war phases defined subsequently in this section.

The belief updating procedure aims to minimize the forecast error, which implies that the mean dynamics trajectory should be near points in the belief space where $u = \mathrm{E}[u_{in}] = 0$. The beliefs satisfying the level curve $u = 0$ are given as

$$\gamma_0 = \frac{2bc - (2a - c)\gamma_1}{2b - \gamma_1}. \tag{1.24}$$

In Figure 1.7, we plot the beliefs satisfying the above equation (1.24) together with the mean dynamics trajectory of beliefs derived from (2.13). From the figure we see that the mean forecast error stays positive during the escape until the cooperative equilibrium is reached, then its value remains for a while near zero and afterwards becomes negative when the belief trajectory crosses the $u = 0$ level curve. Moreover, during the escape the mean forecast error is large in absolute value, which leads to fast belief revision. For the rest of the escape-cooperation-price war cycle, the forecast error is rather small, leading to very slow evolution of beliefs.

The beliefs escape from the SCE when they happen to be outside of the Nash equilibrium region of attraction. Intuitively, this may happen when a particular sequence of shocks induces both firms to contract the supply simultaneously, subsequently leading to an increase of the mean actual price $y = \mathrm{E}[y_n]$. Since firms try to minimize the mean forecast error $u = \mathrm{E}[u_{in}] = \mathrm{E}[y_{in} - \hat{\mathrm{E}}[y_{in}|y^{n-1}]]$, the increase of the mean actual price leads firms to update their beliefs accordingly and to increase the mean forecasted price by contracting the supply. The latter leads to a further increase in the mean expected price. This produces a period of upward price spiral and increasingly cooperative behavior.

After a period of cooperation the gap between the mean actual price and the mean forecasted price approaches zero and the beliefs are in a quasi-equilibrium for a while. Firms perceive this situation as an equilibrium state, as their beliefs are approximately validated by the observations. This period can be characterized as the *cooperative phase*.

However, the perception of the equilibrium state is destroyed when the mean forecast error becomes negative. At this moment the mean actual price falls below the mean forecasted price and the firms start to update their beliefs aggressively in order to reduce the mean price forecast. They achieve this by expanding supply. The expansion of supply implies further decrease of the average actual price, and the updating cycle continues to chase the data. This leads the industry into a period of downward price spiral which moves the beliefs back toward the SCE. Since at the SCE the mean forecast error is zero while during the price war it is negative, there is a point where the absolute value of u stops declining ($\dot{u}/u$ becomes positive) and the belief system is pushed into the attraction region of the Nash equilibrium.

The periods of increasing and decreasing price spirals can be characterized as the *cooperative* and the *price war* phases respectively. To be more specific, we define the *cooperative phase* as the period of increasing price spiral together with the time spent in the neighborhood of the cooperative equilibrium and the *price war phase* as the period of decreasing price spiral till the time when the belief system appears in the attraction region of the NE. Here, we can find some similarities to the Green and Porter model where it is the unanticipated demand shocks that cause price wars when the price falls below the predefined trigger price. In our context the mean forecast price serves as the trigger price, and the unanticipated fall of the mean actual price below the mean forecasted price causes the price war. We thus find that the demand shock is not necessarily the sole source causing price wars, and we could generate the wars even without a demand shock.

□ **Numerical Analysis of Durations.** Our model allows us to investigate the dependence of the durations of cooperative and price war phases on the model parameters. We use mean dynamics ODE (2.13) to calculate belief thresholds that trigger respectively a price war and reversion to the NE. Therefore, for convenience, durations from the real-time dynamical system (2.7) are expressed in continuous time. For the purposes of duration calculations we adopt the following definitions.

Definition 2. *The duration of the cooperation is defined as* $\tau^c = t_2 - t_1$ *where* $t_1 = \inf\{t | \|\gamma_t - \bar{\gamma}\| > \|\gamma^* - \bar{\gamma}\|\}$ *and* $t_2 = \inf\{t | u_t < 0\}$.

The Cooperative phase is defined as the period between the time when beliefs escape the NE for the first time and the time the mean forecast error turns negative for the first time, triggering a price war.

Definition 3. *The duration of the price war is defined as* $\tau^{pw} = t_3 - t_2$ *where* $t_2 = \inf\{t | u_t < 0\}$ *and* $t_3 = \inf\{t | u_t = \min u_s\}$.

The price war is a period when the forecast error is negative and decreasing, forcing the firms to increase output so that the actual price decreases, following the forecasts. The start of the price war phase coincides with the end of cooperation.

We turn now to the numerical results. We depict the duration of the cooperation and the price war phases in Figure 1.8. Both durations are increasing in the Nash equilibrium output $\bar{q}$ and decreasing in volatility σ_2^2 of the firm-specific shock. A higher $\bar{q}$ increases the losses that firms incur when the cooperation breaks down, as the profit difference between cooperative

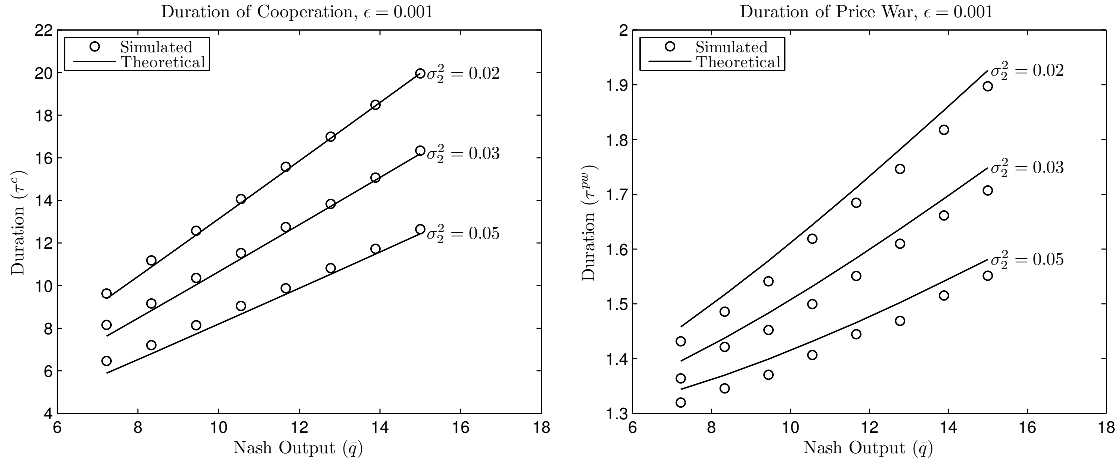


FIGURE 1.8: Durations of cooperation and price war phases as a function of Nash output $\bar{q}$ for a different values of the firm-specific shock volatility σ_2^2

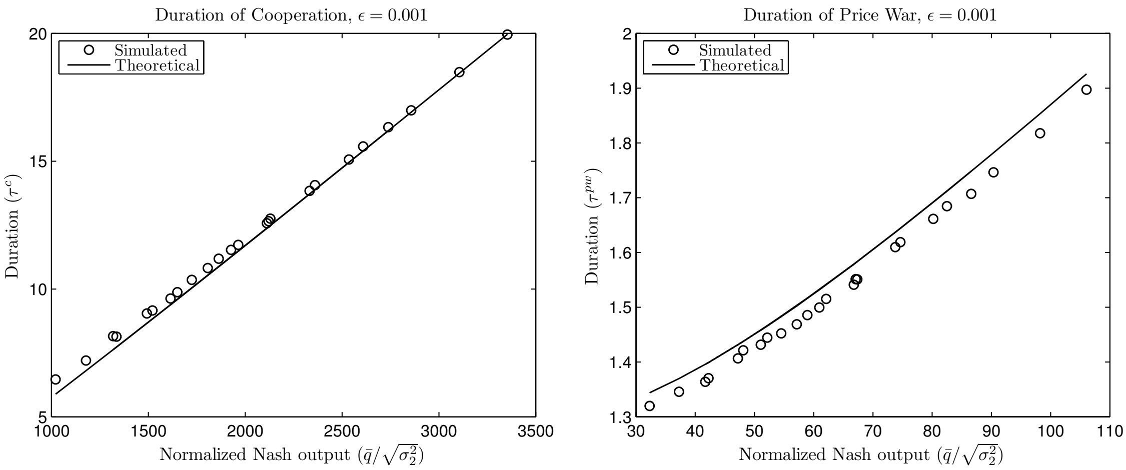


FIGURE 1.9: Durations of cooperation and price war phases as a function of normalized Nash output, $\bar{q}/\sqrt{\sigma_2^2}$.

and Nash equilibria is proportional to $\bar{q}^2$. Therefore, in industries with a higher $\bar{q}$, it is harder to break down a cartel. In industries with a higher σ_2^2, the firms' actions are less correlated. The latter, together with profit maximization, implies a higher output and increases the likelihood of initiating the price war. Moreover, in industries with higher firm-specific shock volatility σ_2^2, the cooperation could only be achieved for sufficiently high levels of the Nash equilibrium output $\bar{q}$. This is the other side of the difficulty of learning to cooperate as evidenced by mean expected exit time discussed above.

In Figure 1.9 and Figure 1.10 , we depict the same curves as in Figure 1.8 and Figure 1.5, respectively, but with normalized Nash output, $\bar{q}^N = \bar{q}/\sqrt{\sigma_2^2}$, on the horizontal axes. We see that the durations turn out to be a function of a single combination of parameters, $\bar{q}/\sqrt{\sigma_2^2}$. The durations of cooperative and price war phases, compared to the mean exit time from the SCE, do

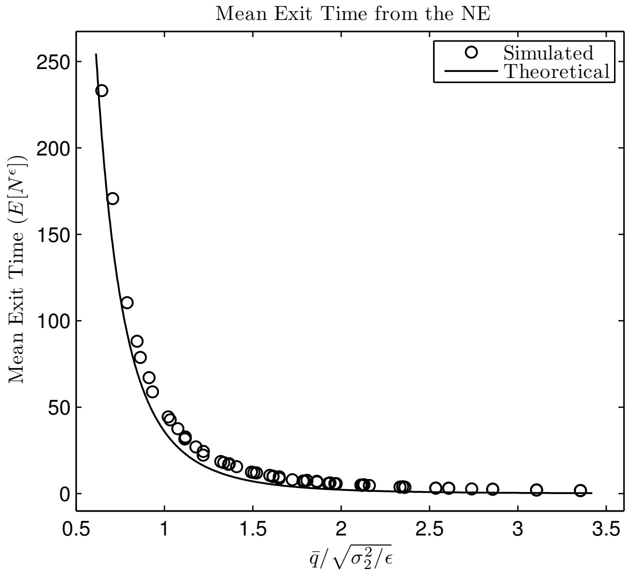


FIGURE 1.10: Mean exit time from the NE as a function of $\bar{q}/\sqrt{\sigma_2^2/\epsilon}$.

not depend on constant gain parameter ϵ since after beliefs escape the SCE, the model dynamics are mainly governed by the mean dynamics. From (1.23), we see that mean exit time depends on the expression $\bar{q}/\sqrt{\sigma_2^2/\epsilon} = \bar{q}^N\sqrt{\epsilon}$. To check this property we depict all the simulation results in both panels of Figure 1.5 with $\bar{q}^N\sqrt{\epsilon}$ on the horizontal axis. Indeed, we see that all the average escape times from real-time dynamic simulations in Figure 1.5 are well approximated by the inverse of the 4^{th} power of $\bar{q}^N\sqrt{\epsilon}$ (solid line).

The above findings are summarized in the following result. The result contains testable hypotheses about the dependence of durations of cooperative and price war phases on the structural parameters of the model.

PROPOSITION 3. *Durations of cooperative and price war phases are increasing functions of the normalized Nash output, $\bar{q}^N = \bar{q}/\sqrt{\sigma_2^2}$. These durations are thus increasing with the industry output $\bar{q}$ and decreasing with the standard deviation of the idiosyncratic shock.*

The Duration of the competitive phase is inversely proportional to the 4^{th} power of $\bar{q}^N\sqrt{\epsilon}$, thus decreasing with the industry output $\bar{q}$ and increasing with the standard deviation of the idiosyncratic shock. The faster the speed with which the firms update their beliefs, the lower is the industry ability to maintain a competitive regime.[15]

As shown in Appendix B, the largest eigenvalue of $\bar{P}$ is approximately equal to $\bar{q}^2/\sigma_2^2 = \left(\bar{q}^N\right)^2$.

[15] Sannikov and Skrzypacz (2007) show when information arrives continuously, under the same settings as in Green and Porter (1984), collusion is impossible. However, under speed of updating we do not mean the frequency of updating but the extent to which new information is incorporated into belief updating. In our model this is measured by ϵ, the constant gain parameter.

Due to the approximate one-dimensionality of the model, this eigenvalue is the most important for both the real-time dynamics (2.7) and the mean dynamics (1.18). Therefore, the evolution of our model is mainly a function of two parameters, normalized Nash output and (for the expected escape to cooperation time) the constant gain, which allows a very parsimonious description of the model dynamics.

1.6. Conclusion

■ In this paper, we study the imperfect monitoring model of duopolistic industry in the sense of Green and Porter (1984) but with unknown structural parameters. Firms build their perceptions about the economy, form beliefs about unknown parameters, construct an estimation model and update their beliefs when new price signals are revealed by the market. We show that when firms take into account the possibility of structural changes in the estimation of parameters, the beliefs of firms may become correlated and this may generate fast escapes from the Nash equilibrium (NE) towards a collusive equilibrium (CE) and then a slow return to the Nash equilibrium. We show that the movement towards the collusive equilibrium constitutes collusive behavior, whereas movement back to the Nash constitutes a *price war*. The fast and slow dynamics are explained by the behavior of the mean forecast error.

The Model's predictions lead to testable hypotheses about the dependence of the likelihood of reversion to cooperation, and durations of the cooperative and price war phases on the structural parameters. We expect that industries with the same normalized output have the same durations of cooperative and price war phases as well as the same mean exit time from the NE. Moreover, industries with a higher normalized output possess longer durations of cooperative and price war phases, but shorter mean exit times from the NE. Therefore, normalized output can serve as a measure to rank industries with respect to the duration of cooperative and price war phases as well as mean exit times from the NE. In addition, we find that it is easier to form a cartel in industries where firms are induced to analyze more common information.

Our results show that a cartel can break down not only due to exogenous factors as in Green and Porter (1984) or Slade (1989) but also due to the endogenous dynamics involved with the uncertainty introduced through the unknown structural parameters. Finally, we show that firms do not have to *a priori* commit to some trigger strategies, and even though playing Nash equilibrium strategies all the time, firms can implicitly coordinate on the length of the duration of the punishment period and the price trigger that triggers cooperation when they update beliefs about the unknown parameters of the inverse demand function.

Appendix A

□ **E-stability of SCE.** The SCE is E-stable when the Jacobian of the belief system (1.18) has all eigenvalues with negative real parts. Define $\theta = (\gamma_i', \gamma_j', \text{vec}(P_i)', \text{vec}(P_j)')'$, then the Jacobian of the system (1.18) takes the following form:

$$J = \begin{bmatrix} \frac{\bar{P}_i \partial \bar{g}_i(\bar{\gamma})}{\partial \gamma_i} & \frac{\bar{P}_i \partial \bar{g}_i(\bar{\gamma})}{\partial \gamma_j} & 0 & 0 \\ \frac{\bar{P}_j \partial \bar{g}_j(\bar{\gamma})}{\partial \gamma_i} & \frac{\bar{P}_j \partial \bar{g}_j(\bar{\gamma})}{\partial \gamma_j} & 0 & 0 \\ -\text{vec}\left(\frac{\bar{P}_i \partial M(\bar{\gamma}_i)\bar{P}_i}{\partial \gamma_i}\right) & 0 & -\text{vec}\left(\frac{\partial \bar{P}_i M(\bar{\gamma}_i)\bar{P}_i}{\partial P_i}\right) & 0 \\ 0 & -\text{vec}\left(\frac{\bar{P}_j \partial M(\bar{\gamma}_j)\bar{P}_j}{\partial \gamma_j}\right) & 0 & -\text{vec}\left(\frac{\partial \bar{P}_j M(\bar{\gamma}_j)\bar{P}_j}{\partial P_j}\right) \end{bmatrix}.$$

Since PMP is a quadratic matrix in P, the sufficient condition for E-stability is that the following matrix has eigenvalues with negative real parts:

$$\bar{h}_\gamma = \begin{bmatrix} \frac{\bar{P}_i \partial \bar{g}_i(\bar{\gamma})}{\partial \gamma_i} & \frac{\bar{P}_i \partial \bar{g}_i(\bar{\gamma})}{\partial \gamma_j} \\ \frac{\bar{P}_j \partial \bar{g}_j(\bar{\gamma})}{\partial \gamma_i} & \frac{\bar{P}_j \partial \bar{g}_j(\bar{\gamma})}{\partial \gamma_j} \end{bmatrix}. \tag{1.25}$$

It is straightforward to show that the matrix is of the form

$$\bar{h}_\gamma = \begin{bmatrix} -1 & 0 & -1/2 & -\bar{q} \\ 0 & -1 & 0 & 0 \\ -1/2 & -\bar{q} & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{1.26}$$

All the eigenvalues of the matrix (1.26) (-1.5, -1, -1, -0.5) have negative real parts. Thus, the SCE is E-Stable.

Appendix B

□ **Derivation of the Diffusion Equation for the Belief System.** In addition to the mean dynamics approximation (1.18), it is possible to approximate the real-time dynamics under adaptive learning by a continuous time diffusion process as follows. Stacking together the system of stochastic difference equations (2.7) for both firms i and j, we can define the following system:

$$\theta_{n+1} = \theta_n + \epsilon \mathcal{H}(\theta_n, \zeta_n), \tag{1.27}$$

where $\theta_n = (\gamma_{in}^T, \gamma_{jn}^T, \text{vech}(P_{in})^T, \text{vech}(P_{jn})^T)^T$ is a vector in which we stack both firms' beliefs and current estimates of the variance-covariance matrices, and $\zeta_n = (x_{in}^T, x_{jn}^T, \omega_{in}, \omega_{jn})^T$ is the state vector.[16]

Evans and Honkapohja (2001) shows that for $\epsilon \to 0$, the time path of the system can be approximated by the following diffusion:

$$d\theta_t = h_\theta'(\theta(t,\theta_0))(\theta_t - \theta(t,\theta_0))dt + \sqrt{\epsilon}\Sigma(\theta(t,\theta_0))^{1/2} d\hat{W}_t, \tag{1.28}$$

where $\theta(t,\theta_0)$ is the solution of the mean dynamics ODE with the initial condition θ_0, $h_\theta'(\cdot)$ is the Jacobian of $\text{E}[\mathcal{H}(\theta_n,\zeta_n)]$, $\hat{W}_t$ is a multi-dimensional Wiener process with $\dim(\hat{W}_t) = \dim(\theta)$ and $\Sigma = \text{cov}[\mathcal{H}(\theta_n,\zeta_n)]$.

[16] The operator vech maps the lower triangular part of a symmetric $n \times n$ matrix into a vector by fully stacking the first column, then the second column from the second element down, and so on till a single (n,n) element of a matrix is added to the vector.

Since the SCE is E-stable (see Section 2.6), following Evans and Honkapohja (2001) it suffices to analyze the time path of (1.28) around the SCE $\bar{\theta}$ which is a stable limit point of the mean dynamics. In addition, one can neglect the dynamics of elements in P_{in} and P_{jn} since they tend to change much more slowly than the beliefs γ and remain almost unchanged before the escape. Considering only the dynamics in beliefs γ, the approximating diffusion becomes[17]

$$d\varphi_t = \bar{h}_\gamma \varphi_t dt + \sqrt{\epsilon}\bar{\Sigma}^{1/2} dW_t, \tag{1.29}$$

where $\varphi_t = \gamma_t - \bar{\gamma}$, $\gamma_t = \left(\gamma_{in}^T, \gamma_{jn}^T\right)^T$, W_t is a multi-dimensional Wiener process with $\dim(W_t) = \dim(\varphi_t)$, $\bar{h}_\gamma = h_\gamma^T(\bar{\theta})$, and $\bar{\Sigma} = \text{cov}[\mathcal{H}_\gamma(\bar{\theta}, \zeta_n)]$.

The diffusion (1.29) is four-dimensional. Following Kolyuzhnov et al. (2014), we can transform it into a one-dimensional diffusion. By pre-multiplying the above equation by v^T and denoting $\hat{\varphi}_t = v^T \varphi_t$, we obtain

$$d\hat{\varphi}_t = -\hat{\varphi}_t dt + \sqrt{\epsilon \lambda_{\bar{\Sigma}}} dz_t, \tag{1.30}$$

where $\lambda_{\bar{\Sigma}}$ and $v = \hat{v}/\|\hat{v}\|$ are respectively the dominant eigenvalue and the standardized dominant eigenvector of $\bar{\Sigma}$, $\hat{v} = (\bar{v}^T, \bar{v}^T)^T$; z_t is a one-dimensional Wiener process. A final transformation involves ignoring the term $-\hat{\varphi}_t$ and setting it to zero; thus, the original real-time learning dynamics is approximated by a one-dimensional Brownian motion.

This approximation is very useful because a very simple formula allows us to derive the time until the one-dimensional process (1.30) leaves any interval of the real line. For the interval $[-d; +d]$, assuming that $z(0) = 0$, the expected time is given by the following formula: $E[\tau^\varepsilon] = \frac{d^2}{\epsilon \lambda_{\bar{\Sigma}}}$ (see Karatzas and Sherve (1991, Eq. 5.62, p. 345)). The boundary of the interval d equals $\hat{\varphi}^* = \delta^* v^T \hat{v} = \sqrt{2}\delta^* \|\bar{v}\|$. Finally, notice that to translate continuous time units, in which τ^ε is expressed, into the number of periods, we need to divide τ^ε by ϵ.

Taking all of the above into account, the expected number of periods until an escape towards cooperation happens is given by the following formula:

$$\begin{aligned} E[N^\epsilon] &= \frac{\hat{\varphi}^{*2}}{\lambda_{\bar{\Sigma}}} \frac{1}{\epsilon^2} = \left(\frac{\delta^*}{b}\right)^2 \frac{1}{\epsilon^2} \\ &= 36 \frac{\sigma_2^4}{\bar{q}^4} \frac{1}{\epsilon^2}. \end{aligned}$$

The formula uses the value of $\lambda_{\bar{\Sigma}} = 2b^2\sigma_2^2\bar{\lambda}_1$, where $\bar{\lambda}_1 \approx \|\bar{v}\|^2/\sigma_2^2$ is the dominant eigenvalue of $\bar{P}$.

Thus, the final dependence of $E[N^\epsilon]$ on the model parameters is given as

$$\mathrm{E}[N^\epsilon] \sim \left(\frac{\bar{q}}{\sqrt{\sigma_2^2/\epsilon}}\right)^{-4}. \tag{1.31}$$

The inverse dependance of the mean exit time on $\overline{q}/\sqrt{\sigma_2^2}$ can be justified by the following observation:

$$\frac{\overline{q}}{\sqrt{\sigma_2^2}} = \sqrt{(\bar{\lambda}_1 - 1)(1 - \bar{\lambda}_2)}, \tag{1.32}$$

[17] See also Sargent and Williams (2005), pp. 376-377.

where $\bar{\lambda}_1$ and $\bar{\lambda}_2$ are respectively the largest and smallest eigenvalues of $\bar{P}$. Indeed, from (1.32) we see that $\bar{\lambda}_1 \geq 1$ and $0 \leq \bar{\lambda}_2 < 1$.

We can notice that an increase of $\overline{q}/\sqrt{\sigma_2^2}$ implies an increase of the ratio $\bar{\lambda}_1/\bar{\lambda}_2$. The latter implies acceleration of belief generation along the direction of the dominant vector of $\bar{P}$ and therefore, on average, less time is needed to exit the attraction region. This explains why mean exit time depends inversely on $\overline{q}/\sqrt{\sigma_2^2}$.

References

Abreu, D., D. Pearce, and E. Stacchetti, "Optimal cartel equilibria with imperfect monitoring," *Journal of Economic Theory*, 1986, *39* (1), 251–269.

Balvers, R. and T. Cosimano, "Actively Learning about Demand and the Dynamics of Price Adjustment," *Economic Journal*, September 1990, *100* (402), 882–98.

Chakrabarti, S., "Collusive equilibrium in Cournot oligopolies with unknown costs," *International Economic Review*, 2010, *51* (4), 1209–1238.

Cho, I., N. Williams, and T. Sargent, "Escaping Nash inflation," *Review of Economic Studies*, 2002, *69* (1), 1–40.

Ellison, G., "Theories of cartel stability and the joint executive committee," *The Rand Journal of economics*, 1994, *25* (1), 37–57.

Ellison, M. and A. Scott, "Learning and price volatility in duopoly models of resource depletion," *Journal of Monetary Economics*, 2013, *60* (7), 806–820.

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

__ , __ , and N. Williams, "Generalized stochastic gradient learning," *International Economic Review*, 2010, *51* (1), 237–262.

Friedman, J., "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies*, 1971, *38* (113), 1–12.

__ , *Market Structure and Competition Policy*, Cambridge University Press, 2000.

Fudenberg, D. and D. K. Levine, "Self-Confirming Equilibrium," *Econometrica*, 1993, *61* (3), 523–545.

— **and** — , “Self-confirming equilibrium and the Lucas critique,” *Journal of Economic Theory*, November 2009, *144* (6), 2354–2371.

— **and J. Tirole**, “A Signal-Jamming Theory of Predation,” *RAND Journal of Economics*, 1986, *17*, 366–376.

Green, E. and R. Porter, “Noncooperative collusion under imperfect price information,” *Econometrica: Journal of the Econometric Society*, 1984, *52* (1), 87–100.

Grossman, S.J, R. E Kihlstrom, and L. J Mirman, “A Bayesian Approach to the Production of Information and Learning by Doing,” *Review of Economic Studies*, October 1977, *44* (3), 533–47.

Kaneko, M., “Some remarks on the folk theorem in game theory,” *Mathematical Social Sciences*, October 1982, *3* (3), 281–290.

Karatzas, I. and S. E. Sherve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, 1991.

Kasa, K., “Learning, Large Deviations, And Recurrent Currency Crises,” *International Economic Review*, 2004, *45* (1), 141–173.

Keller, G. and S. Rady, “Optimal Experimentation in a Changing Environment,” *The Review of Economic Studies*, 1999, *66* (3), 475–507.

Kolyuzhnov, D., A. Bogomolova, and S. Slobodyan, “Escape dynamics: A continuous-time approximation,” *Journal of Economic Dynamics and Control*, 2014, *38* (0), 161–183.

Kreps, D., “Anticipated utility and dynamic choice,” *Frontiers of Research in Economic Theory*, 1998, pp. 1983–1977.

Mirman, L., L. Samuelson, and A. Urbano, “Duopoly signal jamming,” *Economic Theory*, 1993, *3*, 129–149.

Riordan, M., “Imperfect Information and Dynamic Conjectural Variations,” *RAND Journal of Economics*, 1985, *16*, 41–50.

Sannikov, Y. and A. Skrzypacz, “Impossibility of Collusion under Imperfect Monitoring with Flexible Production,” *The American Economic Review*, 2007, *97*, 1794–1823.

Sargent, T., *The Conquest of American Inflation*, Princeton: Princenton University Press, 1999.

__ **and S. Williams**, “Impacts of Priors on Convergence and Escapes from Nash Inflation,” *American Economic Review*, 2005, *8*, 360–391.

__ **, N. Williams, and T. Zha**, “The Conquest of South American Inflation,” *Journal of Political Economy*, 2009, *117* (2), 211–256.

Slade, M., “Strategic pricing models and interpretation of price-war data,” *European Economic Review*, 1990, *34* (2-3), 524–537.

Slade, M. E., “Price Wars in Price-Setting Supergames,” *Economica*, 1989, *56* (223), 295–310.

Stigler, G., “A theory of oligopoly,” *The Journal of Political Economy*, 1964, *72* (1).

Wieland, V., “Monetary Policy, Parameter Uncertainty and Optimal Learning,” *Journal of Monetary Economics*, 2000, *46*, 199–228.

Williams, N., “Escape dynamics in learning models.” PhD dissertation, University of Chicago 2001.

__ , “Adaptive learning and business cycles,” 2003. Princeton University, Manuscript.

Chapter 2

Free entry and social efficiency under unknown demand parameters*

In this article, I examine free entry in homogeneous product markets and its social efficiency. Previous research has shown that under a Cournot oligopoly with fixed setup costs, the free entry equilibrium always delivers excessive entry in homogeneous product markets. In contrast, I demonstrate in this article that free entry along with excessive entry might also lead to a socially insufficient number of firms when a demand parameter uncertainty is considered. My findings support the validity of the traditional wisdom in industrial organization that free entry is desirable for social efficiency and call for revision of restrictive entry regulation practices which have been based on previous research findings.

Keywords: Free Entry, Welfare, Collusion, Beliefs, Learning, Self-Confirming Equilibrium

JEL Classification: D60, D83, D43, L13, L40, L51

*An earlier version of this paper has been published in Janjgava, B. (2013) Free entry and social efficiency under unknown demand parameters, CERGE-EI Working Paper Series, 2013, No. 495. I am grateful to Sergey Slobodyan for helpful discussions and comments at various stages of this research. I would like also to thank Krešimir Žigic and Fabio Michelucci for valuable comments and suggestions. All remaining errors are the responsibility of the author. The financial support of the Czech Science Foundation project No. P402/12/G097 DYME Dynamic Models in Economics is acknowledged.

2.1. Introduction

■ In general, free entry is considered desirable for a society from a social welfare point of view and has been traditional wisdom among economic professions. However, many economists have challenged this view and have shown that in homogeneous product markets with imperfect competition, when firms pay a fixed set-up cost upon entry, the free entry equilibrium number of firms exceeds the socially optimum number of firms, known as the excess entry theorem (see Weizsaker (1980), Perry (1984), Mankiw and Whinston (1986), Suzumura and Kiyono (1987), Berry and Waldfogel (1999), Ohkawa et al. (2005)). In this article, I study the social efficiency of free entry and show that the excess entry theorem might not hold in homogenous product markets when the demand parameter uncertainty is considered.

The most prominent work, which provided the first excess entry theorem results under general settings and provoked intensive research in this area, was Mankiw and Whinston (1986). They find that free entry is not socially desirable due to a so called "business stealing" effect which is present when "... the equilibrium output per firm declines as the number of firms grow". Due to new entrants, incumbent firms are forced to reduce output and entry is more desirable to the new entrants than it is to society, implying excessive entry under free entry.

The excess entry theorem advocates restrictive entry polices; but under restricted entry a fixed number of firms operate in the market, and firms' recognition of their mutual interdependence might propagate collusive pricing behavior. The latter idea goes back to Chamberlin (1929) and is not captured by the excess entry theorem. The higher number of firms might hinder the propagation of such collusive pricing behavior and entry acquires an additional effect, i.e. the "competition" effect. This has an opposite impact on social welfare from the "business stealing" effect and makes entry more desirable for society. Hence, if the "competition" effect is present, the results of the excess entry theorem might not hold, or would depend on the net gain from entry and not only on the "business stealing" effect.

The purpose of my article is to reexamine the social efficiency of free entry under the settings of Mankiw and Whinston (1986), but explicitly modeling for the mutual interdependence of firms and the "competition" effect missing in Mankiw and Whinston's model. The building blocks of my model are the same as in Mankiw and Whinston (1986), but in addition I assume a market with a linear inverse demand curve where demand parameters, a slope and an intercept, are unknown to firms. Firms observe only price and their own output and do not observe their rivals' output.

To make production decisions, firms form beliefs about unknown parameters and unobserved output of rivals. After the production decision is made, a new market price is realized and firms update their beliefs using Bayes' rule. Additionally, beliefs are required to be consistent with the observations. This leads to the notion of the self-confirming equilibrium (SCE) as the solution concept (see Fudenberg and Levine (1993), Sargent (1999), Sargent (2008)).

The SCE is less restrictive than the Nash equilibrium and does not require firms to have correct beliefs off the equilibrium path. Firms are aware of this and are interested in generating more information about the off-equilibrium path by "experimentation" to see if they have appropriate beliefs. The "experimentation" is modeled as random perturbations over a firm's best response to its current beliefs, maximizing the instantaneous expected profit. The "experimentation" is costly in terms of forgone instantaneous expected profit and firms are assumed to keep the variation of random perturbations close to zero.

Given the settings of the model, the SCE coincides with the Nash equilibrium. Williams (2001), in similar settings to those in my model but for the case of a duopoly, show the possibility of escapes from the SCE (Nash equilibrium) towards the cooperative equilibrium.[1] Janjgava and Slobodyan (2011), extending Williams (2001), show that firms might engage in cooperative behavior if their beliefs become interdependent enough to coordinate their actions. The model developed in this article is an extension of the duopoly model, presented in Janjgava and Slobodyan (2011), to the oligopoly case, and I find that their results remain valid in the case of oligopoly as well. In addition, I find the conditions when interdependence, which leads to cooperative behaviors, might arise.

In this article, as firms might be engaged in cooperative behavior, the "competition" effect of entry becomes operative. As far as the "competition" effect enhances social welfare, in contrast to the "business stealing" effect, entry has a higher social value under demand uncertainty than in the case of the perfect information studied by Mankiw and Whinston (1986). I find that for lower fixed-setup costs, the results of the excess entry theorem is altered, and the optimal number of firms under demand uncertainty coincides with the free entry number of firms. Thus, the results of the article highlight the importance of the demand parameter uncertainty for the social desirability of free entry, and find support for the traditional wisdom.

The rest of the article is organized as follows. Section 2.2 presents the model, develops the firms' decision problem, and formation and updating of beliefs. Section 2.3 defines the

[1] Ellison and Scott (2013) extend Williams' model to the case of non-renewable resource markets and study the impact of escape dynamics to the volatility of the market price.

self-confirming equilibrium that coincides with the Cournot-Nash equilibrium of the model and a dominant escape path from it. In Section 2.4, I study the conditions under which implicit coordination towards collusion is impossible and define the collusion frontier. Section 2.5 introduces free entry and considers the social planner's problem under free entry when implicit coordination on collusive outcomes might arise endogenously. Section 2.6 concludes.

2.2. The model setup

■ Consider an oligopolistic industry comprised of N firms producing a single, homogeneous good where explicit collusion is forbidden by the antitrust law. It is assumed that firms possess a homogeneous production technology with a constant marginal cost c and a fixed setup cost K. The industry faces a linear inverse demand schedule:

$$y_n = a - bQ_n, \tag{2.1}$$

where a and b are unknown positive constants, and $Q_n = Q_n^{-i} + q_n^i$ is industry supply; q_n^i stands for a firm $i's$ individual output and Q_n^{-i} is rivals' output and not observed by the firm i.

□ **A firm's decision problem.** Given the linear inverse demand function firm i constructs a linear subjective model of demand

$$y_n^i = {x_{in}}^{\top}\theta_n^i + u_n^i, \tag{2.2}$$

where $x_{in} = (1, q_n^i)^{\top}$; $\theta_n^i = (a - bQ_n^{-i}, -b)^{\top} \in \Theta$ is a vector of unknown parameters ; the error term $u_n^i \sim N(0, \sigma^2)$ captures the possibility of mistakes in constructing the subjective model.

Define history in the period n as $h_n^i = \{y_s, q_s^i\}_{s=1}^{n-1}$. At the beginning of each period n, firm i updates its prior beliefs $\mu_{n-1}^i(\theta_n^i|h_{n-1}^i) = P(\theta_n^i|h_{n-1}^i)$ over Θ, the set of unknown parameters, given the realized price y_{n-1} using the Bayes' rule

$$\mu_n^i(\theta_n^i|h_n^i) \propto P(h_n^i|\theta_n^i, h_{n-1}^i)\mu_{n-1}^i(\theta_n^i|h_{n-1}^i). \tag{2.3}$$

Given the current beliefs $\mu_n^i(\theta_n^i|h_n^i)$, firm i solves the profit maximization problem

$$\pi_n^i = \max_{q_n^i} \int_{\Theta} \left(({x_{in}}^{\top}\theta_n^i + u_n^i - c) q_n^i - K \right) d\mu_n^i(\theta_n^i|h_n^i), \tag{2.4}$$

which yields the following decision rule

$$\hat{q}_n^i = \max\left\{\frac{\hat{\theta}_{0n}^i - c}{-2\hat{\theta}_{1n}^i}, 0\right\}, \tag{2.5}$$

where $\hat{\theta}_n^i = \int_\Theta \theta_n^i d\mu_n^i(\theta_n^i | h_n^i)$.

Firm $i's$ belief μ_n^i is required to be consistent with observations and the solution concept becomes the self-confirming equilibrium (SCE). The SCE is less restrictive than the Nash equilibrium and does not require firms to have correct beliefs off the equilibrium path. Firms are aware of this and are interested in generating more information about the off-equilibrium path by an "experimentation" to see if they have appropriate beliefs.[2]

The "experimentation" is modeled as a random perturbation over firm $i's$ best response to its current beliefs and the final production decision is given as

$$q_n^i = \hat{q}_n^i + \omega_n^i,$$

where ω_n^i is assumed to be an ex post privately observable Gaussian white noise.

The "experimentation" is costly in terms of the forgone expected profit and i is interested in keeping the variation of random perturbation ω_n^i close to zero.

□ **A firm's estimation model.** Firms anticipate that, due to learning, the actual environment is not stationary and thus, the opponents' strategies should not be stationary either. Hence, it is assumed that firms consider unknown parameters θ_n^i as time-varying and adopt a time-varying parameter (TVP) estimation model[3]

$$\begin{aligned} y_n^i &= x_{in}^\top \theta_n^i + u_n^i, \\ \theta_n^i &= \theta_{n-1}^i + \eta_n^i, \end{aligned} \tag{2.6}$$

where $\eta_n^i \sim N(0, V_i)$.

Given the estimation model (2.6), firm i updates their beliefs $\mu_n^i(\theta_n^i | h_n^i)$, represented by a normal distribution

$$\theta_n^i | h_n^i \sim N(\hat{\theta}_n^i, \sigma^2 \hat{P}_n^i),$$

[2] For detailed considerations about the "experimentation" see Fudenberg and Levine (2009).

[3] TVP is commonly used in learning literature in the industrial organization as well as in macroeconomics: see Slade (1989), Balvers and Cosimano (1990), Balvers and Cosimano (1993), Sargent (1999), Cho et al. (2002).

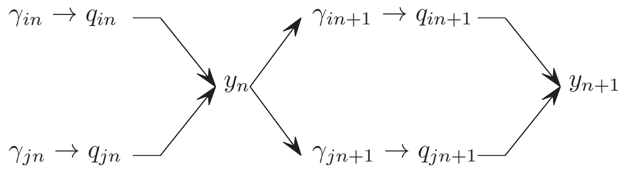


FIGURE 2.1: Timing of the model

recursively using the Kalman filter equations

$$\begin{aligned} \hat{\theta}^i_{n+1} &= \hat{\theta}^i_n + \frac{\hat{P}^i_n}{1 + x^{\top}_{in}\hat{P}^i_n x_{in}} x_{in}(y^i_n - x^{\top}_{in}\hat{\theta}^i_n), \\ \hat{P}^i_{n+1} &= \hat{P}^i_n - \frac{\hat{P}^i_n x_{in} x^{\top}_{in}\hat{P}^i_n}{1 + x^{\top}_{in}\hat{P}^i_n x_{in}} + \frac{1}{\sigma^2}V_i, \end{aligned} \tag{2.7}$$

where $\hat{P}^i_n = \sigma^{-2}\,\text{cov}[\theta^i_n - \hat{\theta}^i_n]$. The Kalman filter, given the normality assumptions, is equivalent to the Bayesian updating rule (3.8).[4]

□ **Timing.** The timing of the decision process is illustrated in Figure 2.1. At the beginning of each period n, given the current beliefs μ^i_n over the set of unknown parameters Θ, firm i makes production decisions q^i_n given by (2.5). After production decisions are made and output is supplied to the market, the market price y_n is realized according to its true data generating process

$$y_n = a - b\sum_{i=1}^{N} q^i_n. \tag{2.8}$$

At the beginning of the next period $n+1$ firm i observes the realized price and updates its prior beliefs μ^i_n over Θ using (2.7) by μ^i_{n+1} and the cycle of production and belief updating process starts over.

2.3. Self-confirming equilibrium

■ Equilibrium in the model is defined by the self-confirming equilibrium (SCE) developed by Fudenberg and Levine (1993) and adopted for adaptive learning literature in macroeconomics by Sargent (1999).

DEFINITION 4. *SCE is a pair of belief and action $(\bar{\theta},\ \bar{q})$ s.t. given the belief $\bar{\theta}$ the action $\bar{q}$ solves a firm's decision problem (2.4), $\bar{q} = \hat{q}^i_n(\bar{\theta})$ and given the action $\bar{q}$ the belief $\bar{\theta}$ is consistent with observations* $\text{E}[x_{in}(y_n - x^{\top}_{in}\bar{\theta})] = 0$.

[4] See for instance Slade (1989), Balvers and Cosimano (1990) for the Kalman filter applications in IO, and for its derivation one can consult Harvey (1989), Ljungqvist and Sargent (2004).

Solving for beliefs in the above orthogonality condition we get

$$\begin{bmatrix} \theta_0^i \\ \theta_1^i \end{bmatrix} = \begin{bmatrix} a - b\mathrm{E}[Q_n^{-i}] + \rho b \mathrm{E}[q_n^i] \\ -b(1+\rho) \end{bmatrix}, \tag{2.9}$$

where $\rho = \mathrm{cov}[\sum_{j\neq i}^{N} q_{jn}, q_n^i]/\mathrm{var}[q_n^i]$ and also can be recognized as a *conjectural variations parameter.*

By definition in the SCE the *conjectural variations parameter* ρ is zero as the SCE action $\bar{q}$ solves a firm's maximization problem (2.4) where the rival's output are taken as given. From the industrial organization literature we know that the *conjectural variations parameter* ρ measures the degree of collusion; $\rho = 0$ corresponds to the Nash equilibrium (NE) whereas $\rho = N-1$ corresponds to the collusive equilibrium (CE). Therefore, the SCE belief and action pair coincides with the NE.

Solving (2.9) for $\rho = 0$ and $\rho = N-1$ provides the following belief and action pairs

$$(\bar{\theta},\ \bar{q}) = \left(\left[\frac{2a+(N-1)c}{N+1}, -b\right]',\ \frac{a-c}{b(N+1)}\right), \tag{2.10}$$

$$(\tilde{\theta},\ \tilde{q}) = \left([a, -bN]',\ \frac{a-c}{2bN}\right), \tag{2.11}$$

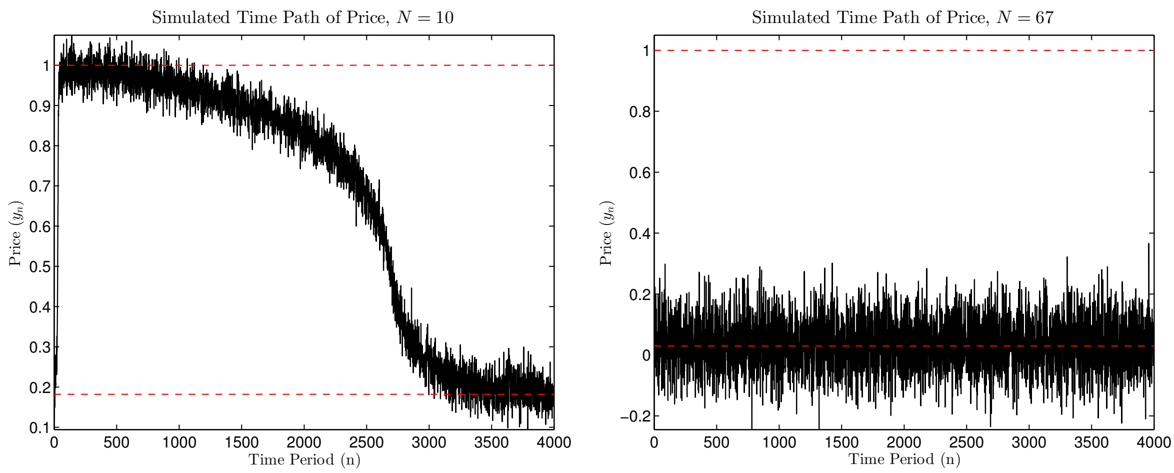
respectively. We can verify that the action $\bar{q}$ at $\rho = 0$ corresponds to the Nash equilibrium output whereas the action $\tilde{q}$ at $\rho = N-1$ corresponds to the collusive equilibrium output where firms maximize joint profits.

□ **Dominant escape path.** Analytical tractability of the model requires the use of a system of ODEs associated with the discreet time belief updating system (2.7). Following Sargent and Williams (2005), the dynamics of $\{\theta_n, P_n\}$ generated by (2.7) converge weakly to the solution of the following system of ODEs, called *mean dynamics*:

$$\begin{aligned} \dot{\theta}^i &= P^i \bar{g}(\theta^i, \ldots, \theta^N), \\ \dot{P}^i &= \sigma^{-2}\bar{V} - P^i M(\theta^i) P^i, \\ i &= 1, \ldots, N, \end{aligned} \tag{2.12}$$

where $\bar{g}(\theta^i, \ldots, \theta^N) = \mathrm{E}[x_{in}(y_n - \overset{\top}{x}_{in}\theta_n^i)]$, $M(\theta^i) = \mathrm{E}[x_{in}\overset{\top}{x}_{in}]$, and $\bar{V} = \sigma^2 M(\bar{\theta})^{-1}$.

We can show that the system of ODEs (2.12) has a unique E-stable steady-state which coincides with the SCE (see Appendix A for derivations). Despite the fact that the SCE is E-stable,



a b

FIGURE 2.2: The simulated time path of price, ($a = 2$, $b = 0.1$, $\sigma_2 = 0.1$, $\epsilon = 1/365$): for a lower number of firms we observe escapes towards the higher CE price (a) whereas for a higher number of firms we do not (b); the red dashed lines shows the SCE/NE and the CE prices.

simulating the belief updating system (2.7) we can observe recurrent escapes from the SCE towards the CE, see Figure 2.2a. At time 0, firms' beliefs constitute the SCE beliefs $\bar{\theta}$ but after some time, when they start to update beliefs, their beliefs escape from the SCE and firms start to coordinate their actions towards the collusive equilibrium price, which for the given parametrization is equal to 1. This behavior depends on the number of firms. For a higher number of firms, beliefs, and consequently the market price, stay in the vicinity of the SCE and the collusive behavior is not observed, see Figure 2.2b.

As firms are homogeneous, imposing symmetry allows us to get rid of the subscript i and concentrate on the mean dynamics of a single firm

$$\begin{aligned} \dot{\theta} &= P\bar{g}(\theta, N), \\ \dot{P} &= \sigma^2 \bar{V} - PM(\theta)P. \end{aligned} \tag{2.13}$$

Following Kolyuzhnov et al. (2014), in the system of ODEs (2.13), if escapes from the SCE happen they occur along the direction of the dominant eigenvector of $\bar{P}$, as $\lambda_1/\lambda_2 \gg 1$ where λ_1 and λ_2 are eigenvalues of $\bar{P}$. [5]

[5] We can perform the following decomposition

$$\bar{P}\bar{g}(\theta, N) = H\Lambda H'\bar{g}(\theta, N) = \lambda_1 c_1 v_1 + \lambda_2 c_2 v_2, \tag{2.14}$$

where Λ is a diagonal matrix formed from the eigenvalues of $\bar{P}$, and the columns of $H = \begin{bmatrix} v_1 & v_2 \end{bmatrix}$ are the corresponding eigenvectors of $\bar{P}$, and $c_i = H_i'\bar{g}(\theta, N)$.

As $\lambda_1/\lambda_2 \gg 1$ the second term in (2.14) becomes negligible and we obtain $\bar{P}\bar{g}(\theta, N) \approx \lambda_1 c_1 v_1$. Thus, the dynamics of the system (2.13) are concentrated along the direction of dominant eigenvector v_1 of $\bar{P}$. For a similar idea one can refer to principal component analysis widely used in the econometrics literature (see for example Stock and

The dominant eigenvector can be approximated by

$$\bar{v} = \begin{bmatrix} \bar{q} \\ -1 \end{bmatrix},$$

which has the direction that coincides with the direction of a line connecting the SCE and the CE beliefs. Thus, the escape path from the SCE towards the CE is given by the following line:

$$\Gamma = \{\theta \mid \theta = \bar{\theta} + \delta \bar{v}, \delta \in \mathbb{R}\}. \tag{2.15}$$

To derive an intuitive explanation of the dominant escape path Γ from the SCE we can look more closely at the orthogonality condition (2.9). The orthogonality condition (2.9) evaluated at the SCE can be expressed as

$$\begin{aligned} \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix} &= \begin{bmatrix} \frac{2a+(N-1)c}{N+1} \\ -b \end{bmatrix} + b\rho \begin{bmatrix} \bar{q} \\ -1 \end{bmatrix} \\ &= \bar{\theta} + b\rho \bar{v}. \end{aligned} \tag{2.16}$$

The expression (2.16) defines a belief evolution in the neighborhood of SCE. This implies that the firms' belief vector θ escapes the SCE whenever the firms' beliefs and consequently firms' actions become interdependent, ρ the conjectural variations parameter becomes positive. Expression (2.16) represents the same relationship as in (2.15) and, together with (2.15) shows that the deviation from the SCE in its neighborhood happens along the direction of $\bar{v}$. As mentioned earlier, the direction of $\bar{v}$ coincides with the direction of a line connecting the SCE and the CE beliefs. Hence, if escape from the SCE happens it happens towards the CE.

The belief updating procedure (2.7) and its stochastic approximation (2.13) drives beliefs in a manner so as to minimize the mean forecast error $u = \mathrm{E}[u_n^i]$. Hence, to understand the properties of the belief updating system one should look at the mean dynamics of the forecast error, namely $\dot{u}/u$.[6] When $\dot{u}/u < 0$, belief updating slows down as the mean forecast error u moves towards zero, whereas when $\dot{u}/u > 0$ belief updating speeds up as the mean forecast error u diverges from zero.

The E-stability of the SCE implies that there is a neighborhood of SCE where $\dot{u}/u < 0$ that

Watson (2002), Bai and Ng (2002), Bernanke and Boivin (2003)).

[6]For a more detailed discussion one can consult Janjgava and Slobodyan (2011).

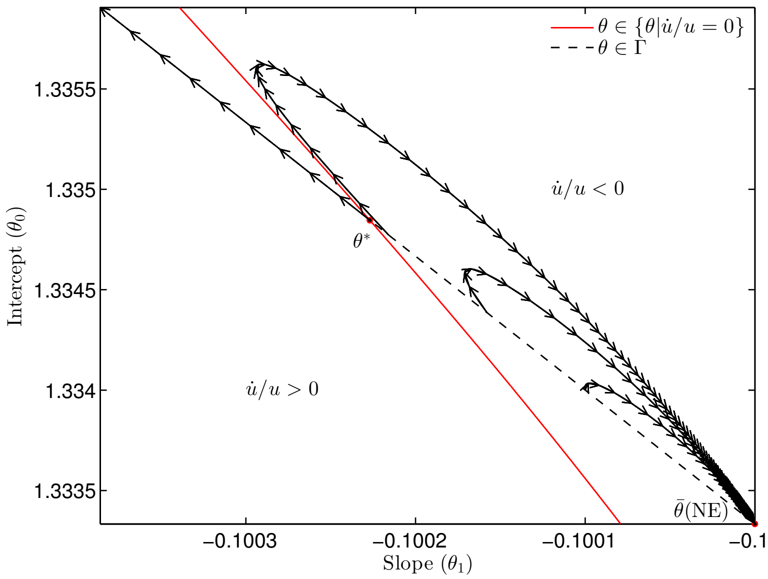


FIGURE 2.3: Phase diagram of beliefs, $a = 2$, $b = 0.1$, $\sigma_2^2 = 0.01$, $N = 2$. Two types of mean belief trajectories are depicted: one with initial beliefs inside the attraction region of SCE along the dominant escape path and another for initial beliefs outside the attraction region of SCE along the dominant escape path. Mean belief trajectories are shown by arrows.

defines an attraction region of the SCE, see Figure 2.3. Whenever the mean forecast error diverges from zero, due to shocks present in the economy, but beliefs still remain in the attraction region of SCE, u starts to converge to zero due to $\dot{u}/u < 0$ and beliefs are attracted back to the SCE. However, when beliefs happen to be outside of the attraction region of SCE due to the updating process, they diverge from the SCE since $\dot{u}/u > 0$ and the mean forecast error u diverges from zero. Thus, beliefs escape the SCE whenever they appear in the region where $\dot{u}/u > 0$.

The propagated dynamics due to belief updating are intuitive enough. The mean forecast error u is defined as the difference between the average price $y = \mathrm{E}[y_n]$ and the expected price $\hat{y}_{in} = {x_{in}}^{\top}\hat{\theta}_{in}$, $u = y - \hat{y}$. When firm i observes a higher average price than he expected, $y - \hat{y} > 0$, he interprets it as a lower demand and reduces output. So, as far as $y - \hat{y} > 0$ firms keep reducing their output and eventually it is closer and closer to the collusive equilibrium (CE) level.

On the other hand, when $y - \hat{y} < 0$, firm i interprets it as a higher demand and increases output. So, as far as $y - \hat{y} < 0$ firms keep increasing their output and eventually it is closer and closer to the Nash equilibrium (NE) output.

2.4. Mutual interdependence and collusion frontier

■ The central question with which the theory of oligopoly is concerned is how firms coordinate with each other when moving towards the collusive outcomes. Cournot thought that cartel agreements may only be maintained by "means of a formal engagement" (Vives, 2000). This view was challenged by Chamberlin (1929). He thought that in markets with a small number of firms the mechanism that elevates prices above competitive levels is the firms' mutual interdependence; the firms' recognition of their interdependence and the futility of cutting prices progressively drives the market price towards the monopoly price.

The model developed by Janjgava and Slobodyan (2011) to study endogenous collusion possibilities captures Chamberlin's reasoning. Janjgava and Slobodyan (2011) find that the belief system (2.7) exhibits escapes from the SCE towards the CE when the degree of firms' mutual interdependence becomes high enough to trigger cooperation. This happens when the interdependence of firms' actions ρ reaches some threshold value ρ^*, an *interdependence threshold*.

Following Janjgava and Slobodyan (2011), we can define the *interdependence threshold* ρ^* as a solution of $\dot{u}/u = 0$. Janjgava and Slobodyan (2011) concentrate on investigating a specific market where the existence of ρ^* is assumed, and do not investigate the properties of $\dot{u}/u = 0$. In this section, I investigate the properties of $\dot{u}/u = 0$ and study the existence of the *interdependence threshold* ρ^*. It is shown that the existence of the *interdependence threshold* ρ^* depends on the market size, the volatility of the firm-specific shock and the number of firms present in the market.

The mean dynamics of the forecast error along the dominant eigenvector $\bar{v}$ direction is given by

$$\frac{\dot{u}}{u} = \left[\frac{(\psi/(N+1)^2+6)}{4((N-1)-\rho)(1+\rho)^3}\right] f(\rho), \tag{2.17}$$

where $f(\rho) = \rho^4 + a_1\rho^3 + a_2\rho^2 + a_3\rho + a_4$ represents the 4-th order *market characteristic polynomial*[7] and

$$\psi = \left(\frac{a-c}{b\sqrt{\sigma_2^2}}\right)^2$$

is a *market characteristic number* and $\sqrt{\psi}$ denotes the normalized competitive market size.[8] In the rest of the article I refer to ψ as a normalized market size.

Escapes from the SCE happens when the belief vector θ hits the $\dot{u}/u = 0$ boundary and $\dot{u}/u$ changes the sign from negative to positive. The $\dot{u}/u = 0$ boundary is determined by the roots of the *market characteristic polynomial* $f(\rho)$. $f(\rho)$ has at least one real root out of its four roots and

[7] One can verify that $\rho = -1$ is not a root of the polynomial $f(\rho)$, $f(-1) = -\frac{b^4(a-c)^2N^2}{(a-c)^2+6b^2(1+N)^2\sigma_2^2} = -b^4N^2\frac{\psi}{(\psi/(N+1)^2+6)}$.

[8] Here, $\frac{a-c}{b}$ is a perfect competition market size where a price equals a marginal cost, y=c.

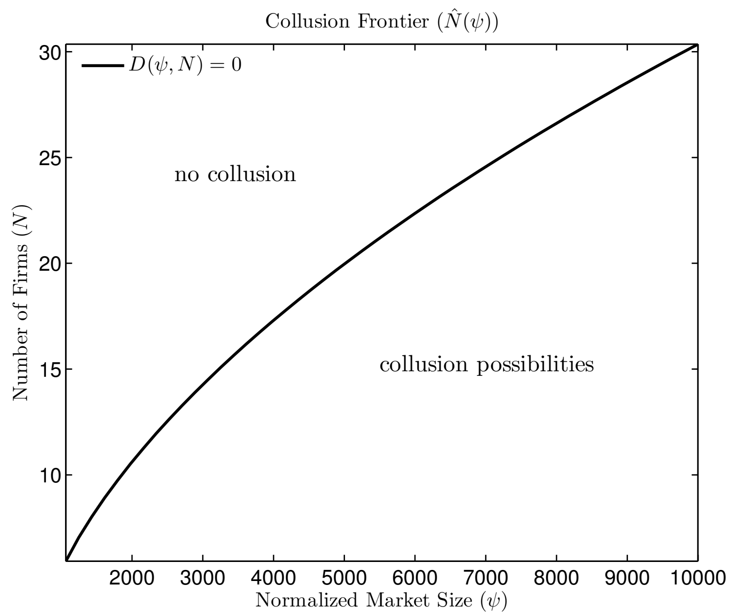


FIGURE 2.4: Collusion frontier $\hat{N}(\psi)$, a solution of $D(N, \psi) = 0$.

when all its roots are real only one of them are negative (see Appendix B). The *interdependence threshold* ρ^* is given by the intersection of $\dot{u}/u = 0$ boundary and the dominant escape path (2.16). So, it is one of the roots of the *market characteristic polynomial* $f(\rho)$.

PROPOSITION 4. *An implicit coordination on collusive outcomes is only possible when there is an interdependence threshold, ρ^* an unstable fixed point of $\dot{u}/u$, such that* $0 < \rho^* \ll 1$. *This holds when all four fixed points of $\dot{u}/u$ are real.*

Proof. See Appendix B. □

Let's suppose that all four fixed points $\rho_1 < \rho_2 < \rho_3 < \rho_4$ of $\dot{u}/u$ are real where $\rho_1 < 0$ and other fixed points are positive. At the SCE both the degree of mutual interdependence ρ and the mean forecast error u are zero. As the SCE is E-stable $\dot{u}/u$ is negative when the degree of mutual interdependence, ρ, is between ρ_1 and ρ_2 and the mean forecast error u converges to zero, leading beliefs to converge to the SCE; whereas $\dot{u}/u$ is positive when $\rho \in (\rho_2, \rho_3)$, and the mean forecast error u diverges from zero, leading beliefs to diverge from the SCE towards the CE. Therefore, ρ_2 is an *interdependence threshold* ρ^*.

In each period, firms update their beliefs, θ, about unknown demand parameters using the same price signals which lead their beliefs to become interdependent. When the level of mutual interdependence ρ reaches the threshold ρ^* firms start to coordinate on collusive outcomes, see Figure 2.3, where θ^* is a belief vector corresponding to the degree of mutual interdependence at $\rho = \rho^*$.

When the number of firms, N, is large, price signals become non-informative due to a large content of firm-specific shocks in the aggregate output promoting belief updating and consequently a possibility of coordination on a collusive outcome to cease. This suggests that there should be some threshold number of firms, $\hat{N}$, such that if $N > \hat{N}$ that no coordination on a collusive outcome is possible.

Following the proof of Proposition 4 in Appendix B the *interdependence threshold* ρ^* exist when the discriminant function of *market characteristic polynomial* $f(\rho)$ is negative, $D(\hat{N}(\psi), \psi) < 0$. So, solving $D(\hat{N}(\psi), \psi) = 0$ for $\hat{N}(\psi)$ gives us a boundary that defines a *collusion frontier*, a maximum number of firms which can support implicit coordination towards the collusion given ψ, see Figure 2.4.

COROLLARY 1. *In a market with a given normalized market size (ψ), there is a maximum number of firms, $(\hat{N}(\psi))$, which can support implicit coordination on collusive outcomes which defines the collusion frontier. In the markets with a market structure such that (ψ, N) is on or above the collusion frontier ($D(\hat{N}(\psi), \psi) \geq 0$) no collusion possibilities may arise. Otherwise there are possibilities of coordination on collusive outcomes with a non-zero probability.*

With the definition of a *collusion frontier*, we are now in a position to analyze dependence of the *interdependence threshold* $\rho^*(\psi, N)$ on a market structure (ψ, N), see Figure 2.5. As was expected, the *interdependence threshold* $\rho^*(\psi, N)$ is decreasing with respect to the normalized market size ψ and increasing with respect to the number of firms N. For example, in the market with $(\psi, N) = (9000, 21)$ the correlation of the firms' actions $\rho^*(\psi, N)$ should be at least about 0.5, firms to become coordinated on collusive outcomes whereas in the market with $(\psi, N) = (9000, 11)$ it only needs to be 0.05 to trigger coordination on a collusive behavior.

PROPOSITION 5. *In markets with a higher normalized market size (ψ) and lower number of firms (N) the coordination towards the collusive outcomes is triggered by a lower degree of correlation among firms'actions $\rho^*(\psi, N)$. Therefore, in these markets it is easier for firms to coordinate and thus, the higher the probability of attaining collusive outcomes.*

The above result is quite intuitive. The higher the normalized market size, the greater the motivation to coordinate towards cooperation due to higher gains from collusion and thus, the lower the degree of mutual interdependence which is needed to attain collusive behavior. However, with a larger number of firms it is difficult to coordinate and the cooperation becomes harder. As a result we can use the *interdependence threshold* $\rho^*(\psi, N)$ to rank industries according

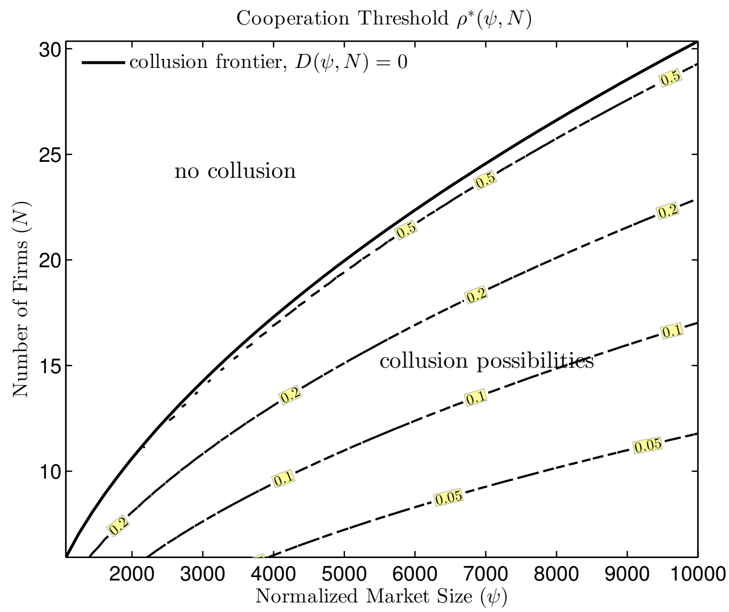


FIGURE 2.5: Interdependence threshold isolines (dashed lines).

to their propensity to collude, given the market structure (ψ, N). For example, in industries with more efficient technologies, the normalized market size is higher and thus it is easier to coordinate towards the collusive outcomes.

2.5. Free entry and social welfare

■ This section deals with the social welfare analysis of free entry. The excess entry theorem shows that in homogeneous markets under free entry equilibrium we have excessive entry. This suggests that policy makers have to restrict entry in these markets and protect incumbents from any new entry. These results are based on the assumption that firms cannot coordinate on collusive outcomes. However, as we saw in the previous section, in homogeneous markets under demand uncertainty firms might implicitly coordinate on collusive outcomes. A priori, the presence of collusion among firms is not desirable for social efficiency since it decreases the social welfare. The restriction of entry as suggested by previous research findings makes it easier for firms to coordinate towards collusion as well as to maintain it, see Proposition 5. Therefore, it is interesting to investigate whether the free entry equilibrium number of firms is still excessive when the possibility of such coordination is considered.

□ **Free entry equilibrium.** Under free entry N^e number of firms are established, determined by a zero profit condition which yields the following expression

$$N^e = \sqrt{\frac{\psi}{(1+\kappa)}} - 1, \tag{2.18}$$

where $\kappa = K/(b\sigma_2^2)$ is a normalized fixed cost.

□ **Social planner's problem.**[9] The social planner's problem is to choose the socially optimal number of firms N^* that maximize a discounted social welfare

$$\begin{aligned} \underset{N}{\text{maximize}} \quad & \int_0^\infty W(\theta, N) e^{-rt}\, dt \\ \text{subject to} \quad & \dot{\theta} = P\bar{g}(\theta, N), \\ & \dot{P} = \sigma^{-2}\bar{V} - PM(\theta)P, \\ & N \leq N^e, \\ & \theta(0) = \bar{\theta},\ P(0) = \bar{P}, \end{aligned} \tag{2.19}$$

where $W(\theta, N)$ is a social welfare function and $r = -\ln\delta$ is a common discount rate. The last constraint serves as the firms' participation constraint and ensures that firms' expected profits are non-negative. When the constraint is binding, free entry provides an insufficient number of firms. However, without subsidy provisions, firms with negative profits are not sustainable and a feasible socially optimal number of firms coincides with the free entry number of firms. Hence, it excludes insufficient entry possibilities.

The social welfare function $W(\theta, N)$ is defined as the sum of consumer and producer surpluses and is given by the following expression

$$W(\theta, N) = \mathrm{E}\left[\int_0^{\Sigma_{i=1}^N q_n^i} y(Q)\, dQ - \left(c\sum_{i=1}^N q_n^i(\theta) + KN\right) | y^{n-1}\right]. \tag{2.20}$$

The entry might have two opposing effects on social welfare. One is the "business stealing" effect of entry, after Mankiw and Whinston (1986), where incumbent firms are forced to reduce output due to new entry and entry is more desirable to new entrants than it is to society. The other one is the "competition" effect that entry increases competition and hinders coordination towards the collusive outcomes. When regulatory authorities solve for the socially optimal

[9] Following the previous literature, only the second best social planner's problem is considered.

number of firms the market structure becomes exogenous. The exogenous market structure stimulates an accommodating pricing behavior and in addition to the "business stealing" effect of entry it makes the "competition" effect of entry operative. However, the "competition" effect of entry is not present in the excess entry theorem. The "business stealing" effect reduces the social welfare whereas the "competition" effect enhances it. Thus, in opposition to the excess entry theorem the net effect of entry on social welfare depends which of these two effects outweighs each other.

Differentiating the expression for the social welfare function (2.20) with respect to the number of firms and evaluating it at the SCE with the free entry number of firms yields

$$\frac{dW}{dN} = \underbrace{(y-c)\frac{\partial q}{\partial N}}_{\text{“business stealing” effect (-)}} + \underbrace{(y-c)\frac{\partial q}{\partial \rho}\frac{\partial \rho}{\partial N}}_{\text{“competition” effect (+)}} . \tag{2.21}$$

The first term is the same "business stealing" effect described by equation (2) in Mankiw and Whinston (1986) whereas the second term is the "competition" effect not present in Mankiw and Whinston (1986). The latter term comes from the fact that the mutual interdependence of firms ρ affects the firms actions and also new entry affects the mutual interdependence of firms ρ.

Let's consider two possible situations $N^* \geq \hat{N}(\psi)$ and $N^* < \hat{N}(\psi)$. When $N^* \geq \hat{N}(\psi)$ no coordination towards collusive behavior happens and beliefs stay in the NE. So, the social planner's dynamic problem (2.19) indeed becomes a static problem and corresponds to the situation considered in Mankiw and Whinston (1986). Thus, whenever $N^* \geq \hat{N}(\psi)$ the socially optimal number of firms N^* is a solution of the same optimization problem as in Mankiw and Whinston (1986) which is given by

$$N^m = \sqrt[3]{\frac{\psi}{\kappa + 1/2}} - 1.$$

When $N^* < \hat{N}(\psi)$ there is a space for coordination among firms towards the collusive outcomes and the social planner's problem is no longer a static but a dynamic problem (2.19). The investigation of the problem is further complicated by the relatively large number of parameters (a, b, c, σ_2, N) that the belief updating system (2.13) depends on. With no loss of generality, to simplify comparative static analysis the following transformation of beliefs are employed:

$$\hat{\theta} = \frac{\theta + [-c, b]'}{b\sqrt{\sigma_2^2}}. \tag{2.22}$$

This delivers the following belief updating system

$$\begin{aligned} \dot{\hat{\theta}} &= P\hat{\hat{g}}(\hat{\theta}, N), \\ \dot{P} &= \sigma^2 \bar{V} - PM(\hat{\theta})P, \end{aligned} \tag{2.23}$$

where

$$\hat{\hat{g}}(\hat{\theta}, N) = \begin{bmatrix} \hat{u} \\ \sqrt{\sigma_2^2}\hat{u}\hat{q} - \sigma_2^2\hat{\theta}_1 \end{bmatrix},$$

$$\hat{u} = \frac{u}{b\sqrt{\sigma_2^2}} = \sqrt{\psi} - (N-1)\hat{q} - \hat{\theta}_0 - \sqrt{\sigma_2^2}\hat{\theta}_1\hat{q},$$

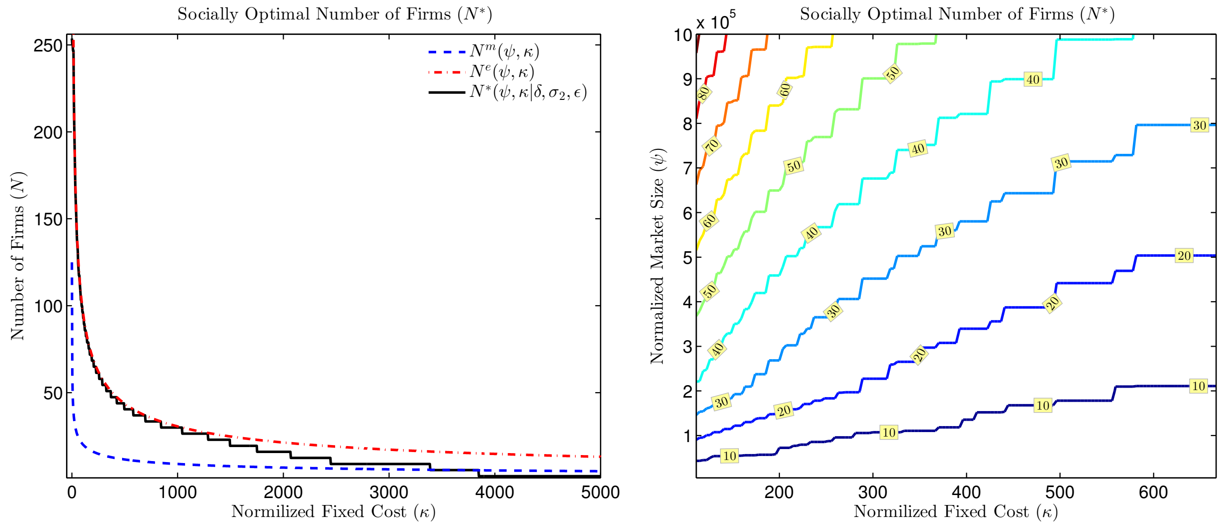
$$\hat{q} = \frac{q}{b\sqrt{\sigma_2^2}} = \frac{\hat{\theta}_0}{2(\hat{\theta}_1 - 1/\sqrt{\sigma_2^2})}.$$

As a results, the dynamics of the transformed beliefs $\hat{\theta}$ depend only on the normalized market size ψ, the standard deviation of firm-specific shock σ_2, and the number of firms N.

The transformation of beliefs (2.22) simplifies the characterization of the parameter dependance of non-cooperative, cooperative and price war phases which are observed during the life cycle of the firms. Firms start in the SCE/NE. After some time $\tau^{NE}(N|\psi,\epsilon)$ when a firm's belief hits the belief threshold $\hat{\theta}^*$ firms start to learn to coordinate on the cooperative behavior and escape from the NE that drives the system towards the CE. The firms maintain the CE beliefs and sustain the collusion for a time period $\tau^{CE}(N|\psi,\sigma_2)$, until the unanticipated fall of mean actual price below the mean forecasted price triggers ($u < 0$), the price war that pushes the system back towards the NE and, after some time, $\tau^{pw}(N|\psi,\sigma_2)$ firms are back again in the NE and the cycle starts again. Let's denote $T(N|\psi,\sigma_2,\epsilon)$ as the time needed for the belief system to escape from the NE towards the CE and return back to it again, thus $T(N|\psi,\sigma_2,\epsilon) = \tau^{NE}(N|\psi,\epsilon) + \tau^{CE}(N|\psi,\sigma_2) + \tau^{pw}(N|\psi,\sigma_2)$. Utilizing this recurrent feature of the belief dynamics after every time period $T(N|\psi,\sigma_2)$, we can rewrite the social planner's objective function in (2.19) as follows

$$V(N|\psi,\kappa,\delta,\sigma_2,\epsilon) = \frac{1}{1-e^{-rT(N|\psi,\sigma_2,\epsilon)}} \int_0^{T(N|\psi,\sigma_2,\epsilon)} \Omega(\hat{\theta}, N|\psi,\kappa,\sigma_2,\epsilon)e^{-rt}\,dt, \tag{2.24}$$

where $\Omega(\hat{\theta}, N|\psi,\kappa,\sigma_2,\epsilon) = \frac{W(\hat{\theta},N)}{b\sigma_2^2}$ is the normalized social welfare function.



a b

FIGURE 2.6: The solution of the social planner's problem and its comparative statics, ($\delta = 0.98$, $\sigma_2 = 0.1$, $\epsilon = 1/365$): (a) depicts the socially optimal number of firms $N^*(\psi,\kappa|\delta,\sigma_2,\epsilon)$ (solid line) and compares it with the socially optimal number of firms $N^m(\psi,\kappa)$ (dashed blue line) when the possibility of coordination towards collusion is ignored, as well as with the free entry equilibrium number of firms $N^e(\psi,\kappa)$ (dash-dot red line). The normalized market size ψ is set at 10^6 and $\hat{N}(\psi) = 348$; (b) depicts the isolines of the socially optimal number of firms $N^*(\psi,\kappa|\delta,\sigma_2,\epsilon)$ and shows its comparative statics properties.

A natural upper bound for the solution is given by the following expression

$$N^* \leq \min\{\hat{N}(\psi), N^e(\psi,\kappa)\}.$$

There is no analytical solution to the problem (2.19) and we have to rely on numerical analysis.

☐ **Numerical results and comparative statics.** Numerical results are obtained with the parameter values of (δ, σ_2, ϵ) set at (0.98, 0.10, 1/365). The numerical solution of the social planner's problem and its properties are depicted in Figure 2.6. Figure 2.6a compares the socially optimal number of firms $N^*(\psi,\kappa|\delta,\sigma_2,\epsilon)$, under demand uncertainty, both with the free entry equilibrium number of firms $N^e(\psi,\kappa)$, and the socially optimal number of firms $N^m(\psi,\kappa)$, under the perfect information of Mankiw and Whinston (1986). We can see that for lower values of the normalized fixed cost κ, the participation constraint (zero profit condition) is binding and in contrast to previous studies, such as Mankiw and Whinston (1986), and subsequent research, the free entry equilibrium delivers an insufficient rather than excessive entry when the "competition effect" of entry is modeled explicitly.

The comparative statics of the optimal solution $N^*(\psi,\kappa|\delta,\sigma_2,\epsilon)$ can be examined simply by plotting its isolines, see Figure 2.6b. From the isolines of the solution we can see that

$N^*(\psi, \kappa|\delta, \sigma_2, \epsilon)$ is an increasing function with respect to the normalized market size ψ, whereas with respect to the normalized fixed cost κ it is decreasing, as can also be seen in Figure 2.6a.

The results are quite intuitive. positive. The higher the normalized market size ψ, the higher the possibility of coordination towards the collusive outcome (see Proposition 5), and controlling for the "competition" effect becomes even more important when assessing the socially optimal number of firms. Thus, the higher the normalized market size ψ, the higher the gain in social welfare from additional entry in hindering possibilities of coordination on a collusive outcome. On the other hand, the higher the normalized market size ψ, less business should be "stolen" with each additional entry. Hence, the socially optimal number of firms is increasing with respect to the normalized market size ψ.

The entry effect on social welfare with regards to the normalized fixed costs are slightly different, as it affects only the "business stealing" effect of entry, but not the possibility of coordination towards collusion. When the normalized fixed cost is high, more business needs to be "stolen" from existing firms for a new firm to operate profitably. Thus, with a high normalized fixed cost, the net marginal effect of entry on social welfare is negative and the free entry number of firms exceeds the socially optimal number of firms.

The results can be summarized in the following proposition:

PROPOSITION 6. *In markets with a low normalized fixed cost (κ) and a high enough normalized market size (ψ), the zero profit condition might be binding and the number of firms under free entry equilibrium might be socially insufficient rather than excessive.*

The results show that the findings of Mankiw and Whinston (1986) are a special case and whether the free entry equilibrium number of firms is excessive or not depends on the market structure and conditions when demand parameter uncertainty is considered. Hence, the traditional wisdom in industrial organization that free entry is desirable for social efficiency is valid in these markets.

2.6. Conclusion

■ In this article, I study the social efficiency of free entry under demand parameter uncertainty. Previous research on free entry and its welfare implications in homogeneous product markets under perfect information shows that the free entry equilibrium delivers excessive entry, known as the excess entry theorem, and suggests restrictive entry regulation policies. I have shown that the introduction of demand parameter uncertainty alters the results of the excess entry theorem and the restrictive entry regulation policies might facilitate collusion possibilities.

When entry is restricted by regulatory authorities, the market structure becomes exogenous. The theory of endogenous market structure has shown that exogenous entry stimulates firms to coordinate their actions towards accommodating pricing behavior to increase prices (Etro (2010)). However, the excess entry theorem does not consider this fact and its predictions are solely based on the "business stealing" effect of entry, where incumbent firms are forced to reduce output due to a new entry, and entry is more desirable to new entrants than it is to society. Under restricted entry a fixed number of firms operate in the market and the abuse of market power and the possibility of accommodating pricing behavior makes the additional effect of entry, the "competition" effect, operative. The "competition" effect reduces market power and hinders the possibility of coordination on a collusive outcome. These two effects of entry impact differently on social welfare, and whether entry is socially efficient depends both on the "business stealing" effect and on the "competition" effect.

In this article, I show that learning the unknown demand parameters when market structure is exogenous leads firms to implicitly coordinate on collusive outcomes and allows the study of the "competition" effect of entry along with the "business stealing" effect. I find that the competition effects of entry outweigh the "business stealing" effects of entry in the markets with a high enough competitive market size. This supports the validity of the traditional wisdom in industrial organization that the free entry is desirable for social efficiency in homogeneous product markets.

Additionally, I derive the *collusion frontier*, which shows the maximum number of firms which might support implicit coordination on collusive outcomes due to demand uncertainty in the homogeneous market given the market size. This can be used to assess the bias of a specific market towards possibilities of implicit coordination on collusive outcomes.

Appendix A

□ **E-stability of the SCE.** To find the steady-state of the system of ODEs (2.12) we have to solve $\dot{\theta}^i = 0$ for all $i \in \{1, ..., N\}$. Since P^i is positive definite matrix, a solution for the steady-state can be obtained through solving $\bar{g}(\theta^i, \ldots, \theta^N) = 0$ for all $i \in \{1, ..., N\}$.

Let, $q^i = \mathrm{E}[q^i_n]$, $y = \mathrm{E}[y_n] = a - b\sum_{i=1}^N q^i$, and $u^i = \mathrm{E}[u^i_n] = y - \theta^i_0 - \theta^i_1 q^i$, then by definition $\bar{g}(\theta^i, \ldots, \theta^N, N)$ is given by the following expression

$$\bar{g}(\theta^i,\dots,\theta^N) = \begin{bmatrix} u^i \\ u^i q^i - (b+\theta_1^i)\sigma_{2i}^2 \end{bmatrix}.$$

Solving $\bar{g}(\theta^i,\dots,\theta^N,N) = 0$ gives

$$u^i = 0, \tag{2.25}$$

$$\theta_1^i = -b. \tag{2.26}$$

Using (2.26) and (2.5) in (2.25) we get

$$\theta_0^i = 2y - c.$$

Therefore, as was expected the steady-state value of the belief vector θ^i are the same among the firms and is given by

$$\theta^i = \left[\frac{2a+(N-1)c}{N+1}, -b\right]',$$

which is indeed the SCE belief $\bar{\theta}$.

Following Evans and Honkapohja (2001), the sufficient condition for E-stability of the system of ODEs (2.12) is that the Jacobian of its belief updating part at the SCE has all eigenvalues with a negative real part. The Jacobian is given by the following matrix

$$\underset{N\times N}{\bar{h}_\theta} = \begin{bmatrix} \frac{\bar{P}_1\partial\bar{g}_1(\bar{\theta})}{\partial\theta_1} & \frac{\bar{P}_1\partial\bar{g}_1(\bar{\theta})}{\partial\theta_2} & \cdots & \frac{\bar{P}_1\partial\bar{g}_1(\bar{\theta})}{\partial\theta_N} \\ \frac{\bar{P}_2\partial\bar{g}_2(\bar{\theta})}{\partial\theta_1} & \frac{\bar{P}_2\partial\bar{g}_2(\bar{\theta})}{\partial\theta_2} & \cdots & \frac{\bar{P}_2\partial\bar{g}_2(\bar{\theta})}{\partial\theta_N} \\ \vdots & \cdots & \ddots & \vdots \\ \frac{\bar{P}_N\partial\bar{g}_N(\bar{\theta})}{\partial\theta_1} & \frac{\bar{P}_N\partial\bar{g}_N(\bar{\theta})}{\partial\theta_2} & \cdots & \frac{\bar{P}_N\partial\bar{g}_N(\bar{\theta})}{\partial\theta_N} \end{bmatrix}, \tag{2.27}$$

where

$$\frac{\bar{P}_i\partial\bar{g}_i(\bar{\theta})}{\partial\theta_i} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \frac{\bar{P}_i\partial\bar{g}_i(\bar{\theta})}{\partial\theta_j} = \begin{bmatrix} -1/2 & \bar{q} \\ 0 & 0 \end{bmatrix} = B.$$

We can rewrite the Jacobian $\bar{h}_\theta$ in a compact way

$$\bar{h}_\theta = (J-I)\otimes B - I,$$

where J is a $N\times N$ matrix of ones and I is a identity matrix with ones on the main diagonal and zeros elsewhere.

It is well known or we can easily verify that J has $N-1$ multiple eigenvalues of 0 and one eigenvalue of N. Let the spectrum of a matrix $J-I$, and B matrices be λ and μ respectively, then we have

$$\lambda = (-1,\dots,-1,N-1),\ \mu = (-1/2, 0).$$

The eigenvalues of the Kronecker product of $J-I$, and B matrices $(J-I)\otimes B$ are

$$\lambda^i\mu^i,\ i = 1,\dots,N, j = 1,\dots,N.$$

$$(1/2,\dots,1/2,-(N-1)/2,0,\dots,0).$$

Finally, the eigenvalues of $\bar{h}_\theta$ are given by

$$\lambda^i\mu^i - 1,\ i = 1,\dots,N,\ j = 1,\dots,N.$$

$$(-1/2,\dots,-1/2,-(N-1)/2-1,-1,\dots,-1).$$

We see that all eigenvalues of the Jacobian $\bar{h}_\theta$ are negative, which proves the E-stability of the system of ODEs (2.12).

Appendix B

☐ **Analyzing the Market Characteristic Polynomial.** This appendix provides proof for Proposition 4. First, let's prove the following two propositions before dealing directly with Proposition 4.

Proposition 7. *The market characteristic polynomial $f(\rho)$ has at least one real root. Therefore, it has either all four roots real or two real and two complex conjugate roots since $f(\rho)$ is a 4-th order polynomial.*

Proof. Since $f(0) < 0$, $f(-\infty) > 0$ and $f(+\infty) > 0$ the market characteristic polynomial can not have all roots complex. Moreover, $f(\rho)$ is 4-th degree polynomial and therefore, it can have either four real roots (one negative and three positive or three negative and one positive) or two real (one positive and one negative) and two complex conjugate roots. ☐

The market characteristic polynomial $f(\rho)$ is associated with the following discriminant function.

$$D(\psi,N) = A_1(N)\psi^5 + A_2(N)\psi^4 + A_3(N)\psi^3 + A_4(N)\psi^2 + A_5(N)\psi + A_6(N). \tag{2.28}$$

Whether $f(\rho)$ has all roots real or two real and two complex conjugate roots depends on the sing of its discriminant:[10]

$D(\psi,N) > 0$	two real and two complex conjugate roots,
$D(\psi,N) = 0$	two simple real and one twofold real roots,
$D(\psi,N) < 0$	four real roots.

Solving $D(\psi,N) = 0$ for N we obtain a threshold boundary $\hat{N}(\psi)$, the number of firms supporting two simple real and one twofold real roots, dividing the parameter space into two subspaces, $D(\psi,N) \geq 0$ and $D(\psi,N) < 0$, Figure 2.7. For $N \geq 2$, $D(\psi,N) < 0$ is possible when $\psi > \underline{\psi} = 1064$. So, the lower bound for ψ s.t $D(\psi,N) < 0$ for $N \geq 2$ is $\underline{\psi} = 1064$.

Proposition 8. *The market characteristic polynomial $f(\rho)$ has only one negative root when all its roots are real.*

[10]One can check that it is not possible to have one simple real and one threefold real roots since ψ is positive.

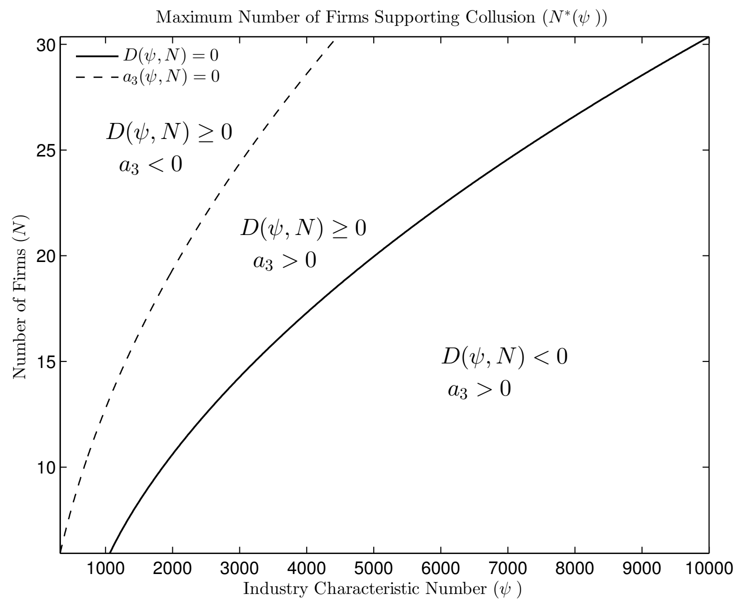


FIGURE 2.7: Shows the maximum number of firms supporting collusion $\hat{N}(\psi)$ given market characteristic number ψ and also that a_3 is positive when all roots of $f(\rho)$ are real.

Proof. According *Descartes' rule of signs*, to have only one negative root it is enough to show that coefficients of $f(\rho)$ $a_2 < 0$ and $a_3 > 0$ when all roots are real, $D(\psi, N) < 0$. We can check that $a_3 > 0$ implies $a_2 < 0$. For $N \geq 2$, $a_3 > 0$ when

$$\psi > \frac{4(N+1)^2(N(N+2)-6)}{(N-1)^2}.$$

From Figure 2.7 we see that when all roots are real ($D(\psi, N) < 0$) $a_3 > 0$. This proves that when all roots are real $f(\rho)$ can only have one negative root. □

Now, we can go back to Proposition 4. From the above propositions we know that we can have only two possibilities: 1) four real roots (one negative and three positive) and 2) two real (one positive and one negative) and two complex conjugate roots.

Case 1: $D(\psi, N) < 0$, four real roots ρ^i, $i = \{1,2,3,4\}$ (in the increasing order)

Since the SCE is E-stable $f(\rho)$ is negative for $\rho \in (\rho_1, \rho_2) \bigcup (\rho_3, \rho_4)$ and positive otherwise. Therefore, when $\rho \in (\rho_2, \rho_3)$, $\dot{u}/u$ becomes positive and belief system escapes from the SCE and we observe cooperative behavior. Thus, ρ_2 serves as an *interdependence threshold* $\rho^*(\psi, N)$. The cooperative behavior to be maintained, the belief updating system should drive ρ to cross the threshold $\rho^*(\psi, N)$ in a finite time. Figure 2.8 depicts that for the range of parameter values, (ψ, N) a mean hitting time to the threshold $\rho^*(\psi, N)$ is finite and therefore the mean exit time from the NE is finite as well. This effectively means that when $D(\psi, N) < 0$ we might observe implicit coordination on collusive outcomes and escape from the NE towards the CE.

Case 2: $D(\psi, N) \geq 0$, two real and two complex conjugate roots

For the parameter values when $D(\psi, N)$ becomes non-negative ρ_2 and ρ_3 becomes complex numbers and we are left only with two real roots ρ_1 and ρ_4. Therefore, we may observe escapes from the NE only when $\rho > \rho_4$. The ρ_4 is the largest positive root of $f(\rho)$ and now the belief system needs more time to exit from the NE than in the case

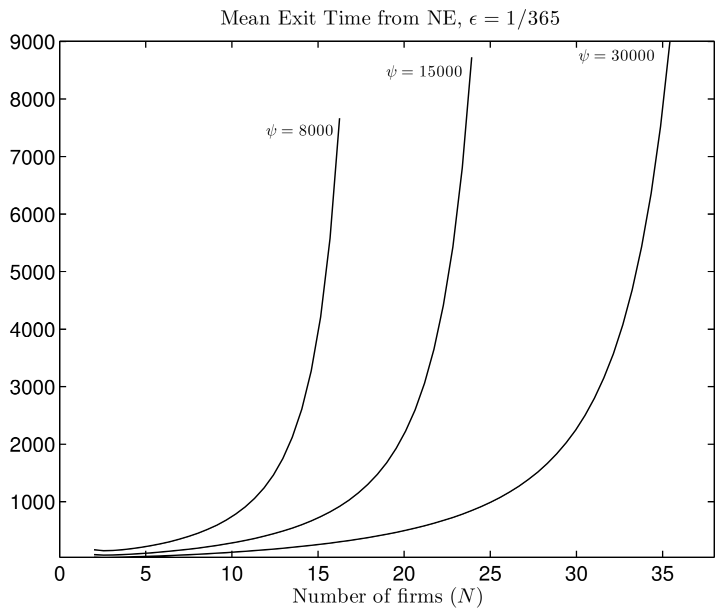


FIGURE 2.8: Mean exit time from NE when $D(\psi, N) < 0$

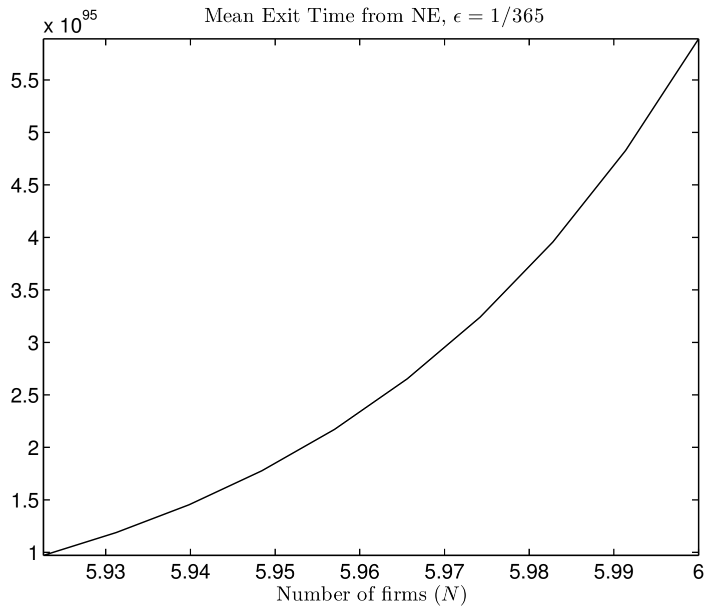


FIGURE 2.9: Mean exit time from NE when $D(\psi, N) \geq 0$.

of $D(\psi, N) < 0$. The bounds for the largest positive root of $f(\rho)$ are given by

$$(N+1)\left(\sqrt{\frac{2N}{N+1}} - 1\right) \leq \rho_4 \leq (N-1).$$

It is a nice observation that bounds does not depend on ψ. Even using the lower bound we can show that the mean exit time from the NE when $D(\psi, N) \geq 0$ is a large number, Figure 2.9.

Therefore, we can conclude that when $D(\psi, N) \geq 0$ we do not observe implicit cooperation among the firms whereas otherwise we might. This effectively means that in an industry with the *characteristic number* ψ a tacit collusion can be observed if, and only if, $D(\psi, N) < 0$, or in other words the number of firms does not exceed $\hat{N}(\psi)$, $N \in [2, \ldots, \hat{N}(\psi))$. This proves Proposition 4.

□ **Mean Hitting Time.** Consider a one-dimensional Ornstein-Uhlenbeck process

$$d\xi = -\xi dt + \sqrt{2} dW,$$

where W is a Wiener process.

Following Ricciardi and Sato (1988), the mean hitting time to a threshold ξ^* for the Ornstein-Uhlenbeck process ξ when it start at zero can be evaluated with the following expression

$$\mathrm{E}[\tau] = \frac{1}{2} \sum_{k=1}^{\infty} \frac{(\sqrt{2}\xi^*)^k}{k!} \Gamma\left(\frac{k}{2}\right), \tag{2.29}$$

where $\tau = \inf\{t | \xi_t = \xi^*\}$ and $\Gamma(\cdot)$ is the gamma function.

Following Kolyuzhnov et al. (2014), we can show that belief dynamics around the SCE are approximated by the following one-dimensional Ornstein-Uhlenbeck process

$$d\hat{\varphi} = -\hat{\varphi} dt + \sqrt{\epsilon \lambda_{\bar{\Sigma}}} dW, \tag{2.30}$$

where $\hat{\varphi} = v'(\theta - \bar{\theta})$, $v = \hat{v} / \| \hat{v} \|$, $\hat{v} = (\bar{v}'_1, \ldots, \bar{v}'_N)$, $\bar{v}'^i = \bar{v}$, and $\lambda_{\bar{\Sigma}} = 2b^2(N-1)\| \bar{v} \|$.

We are interested in the mean hinting time of $\hat{\varphi}$ to the threshold $\hat{\varphi}^*$ which can be assessed using the expression (2.29) with the transformed threshold $\xi^* = \sqrt{2/(\epsilon\lambda_{\bar{\Sigma}})}\hat{\varphi}^*$. The expression for the threshold ξ^* is given by

$$\xi^* = \frac{1}{\sqrt{\epsilon}} \sqrt{\frac{N}{N-1}} \rho^*(\psi, N).$$

References

Bai, Jushan and Serena Ng, "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, January 2002, *70* (1), 191–221.

Balvers, R. and T. Cosimano, "Actively Learning about Demand and the Dynamics of Price Adjustment," *Economic Journal*, September 1990, *100* (402), 882–98.

Balvers, Ronald J. and Thomas F. Cosimano, "Periodic learning about a hidden state variable," *Journal of Economic Dynamics and Control*, 1993, *17* (5-6), 805–827.

Bernanke, Ben S. and Jean Boivin, "Monetary policy in a data-rich environment," *Journal of Monetary Economics*, April 2003, *50* (3), 525–546.

Berry, Steven T. and Joel Waldfogel, "Free Entry and Social Inefficiency in Radio Broadcasting," *RAND Journal of Economics*, Autumn 1999, *30* (3), 397–420.

Chamberlin, E., "Duopoly: Value Where Sellers are Few," *The Quarterly Journal of Economics*, 1929, *44* (1), 63–100.

Cho, I., N. Williams, and T. Sargent, "Escaping Nash inflation," *Review of Economic Studies*, 2002, *69* (1), 1–40.

Ellison, M. and A. Scott, "Learning and price volatility in duopoly models of resource depletion," *Journal of Monetary Economics*, 2013, *60* (7), 806–820.

Etro, Federico, "Endogenous market structures and antitrust policy," *International Review of Economics*, 2010, *57*, 9–45. 10.1007/s12232-009-0088-3.

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

Fudenberg, D. and D. K. Levine, "Self-Confirming Equilibrium," *Econometrica*, 1993, *61* (3), 523–545.

Fudenberg, Drew and David K. Levine, "Learning and Equilibrium," *Annual Review of Economics*, 2009, *1* (1), 385–420.

Harvey, A., *Forecasting, Structural Time Series Models and Kalman Filter*, Cambridge: Cambridge University Press, 1989.

Janjgava, Batlome and Sergey Slobodyan, "Duopoly Competition, Escape Dynamics and Non-cooperative Collusion," CERGE-EI Working Papers wp445, The Center for Economic Research and Graduate Education-Economic Institute, Prague September 2011.

Kolyuzhnov, D., A. Bogomolova, and S. Slobodyan, "Escape dynamics: A continuous-time approximation," *Journal of Economic Dynamics and Control*, 2014, *38* (0), 161–183.

Ljungqvist, Lars and Thomas J Sargent, *Recursive Macroeconomic Theory*, The MIT Press, 2004.

Mankiw, N. Gregory and Michael D. Whinston, "Free Entry and Social Inefficiency," *RAND Journal of Economics*, Spring 1986, *17* (1), 48–58.

Ohkawa, Takao, Makoto Okamura, Noritsugu Nakanishi, and Kazuharu Kiyono, "The Market Selects The Wrong Firms In The Long Run," *International Economic Review*, November 2005, *46* (4), 1143–1165.

Perry, M. K., "Scale Economies, Imperfect Competition, and Public Policy," *Journal of Industrial Economics*, March 1984, *32* (3), 313–33.

Ricciardi, L. M. and S. Sato, "First-passage-time density and moments of the Ornstein-Uhlenbeck process," *Journal of Applied Probability*, 1988, *25*, 43–57.

Sargent, T., *The Conquest of American Inflation*, Princeton: Princenton University Press, 1999.

_ , "Evolution and Intelligent Design," *American Economic Review*, September 2008, *98* (1), 5–37.

_ **and S. Williams**, "Impacts of Priors on Convergence and Escapes from Nash Inflation," *American Economic Review*, 2005, *8*, 360–391.

Slade, M. E., "Price Wars in Price-Setting Supergames," *Economica*, 1989, *56* (223), 295–310.

Stock, J. H. and M. W. Watson, "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 2002, *97* (460), 1167–1179.

Suzumura, K. and K. Kiyono, "Entry Barriers and Economic Welfare," *Review of Economic Studies*, January 1987, *54* (1), 157–67.

Vives, X., *Oligopoly Pricing: Old Ideas and New Tools*, MIT Press, 2000.

Weizsaker, C.C., "A Welfare Analysis of Barriers to Entry," *Bell Journal of Economics*, Autumn 1980, *11* (2), 399–420.

Williams, N., "Escape dynamics in learning models." PhD dissertation, University of Chicago 2001.

Chapter 3

Cooperation in the prisoner's dilemma under belief-based learning

This paper studies a long-run outcome of belief-based learning process in the prisoners' dilemma with anonymous random matching played by a continuum of players. Players have imperfect information about their opponents' play and do not know the payoff matrix of the game. Differing from standard belief-based models, I assume that players hold beliefs not over the likely play of others but over the payoffs they expect from different actions. Players receive public and private signals about payoffs. At the beginning of each round, players update their beliefs using Bayes' rule upon the arrival of new signals. The results of the paper show that for any initial beliefs, only cooperation is sustained in the long-run when players receive perfectly precise public signals.

Keywords: Learning, Cooperation, Self-Confirming Equilibrium, Stochastic Stability Analysis

JEL Classification: D83,D43, L13, L40

3.1. Introduction

Although the Nash equilibrium is the essential and self-enforcing solution concept in game theory, it is very demanding in terms of informational content and does not provide mechanisms to study out-of-equilibrium behavior (Aumann and Brandenburger, 1995). The interest in studying out-of-equilibrium behavior has motivated the development of theory of learning in economics. A survey of recent developments in this field is provided in Fudenberg and Levine (1998), Young (2004), Fudenberg and Levine (2009).

Mostly, learning models are concerned with examining the long-run equilibrium outcome of the learning process when a stage game is repeatedly played by a fixed set of players. These models fall under two broad classifications, namely, *belief-based* and *aspiration-based* learning. In belief-based learning models, a player holds beliefs over opponents' likely play, and her behavior is a best response to her own beliefs. In aspiration-based learning models on other hand, decisions are based on "satisficing", i.e. the "win-stay, lose-shift" principle and instead of solving a expected maximization problem, a player has an *aspiration* level and continues playing a certain strategy as long as the generated payoff exceeds her aspiration level. These models provide different predictions about the long-run outcome of the learning process. In games with a unique strict Nash equilibrium, belief-based models predict the Nash equilibrium as the only possible long-run outcome (see Fudenberg and Levine (1998)). Aspiration-based models, on the other hand, allow players to play strictly dominated strategies (see Karandikar et al. (1998), Kim (1999), Cho and Matsui (2005)).

The, possibility of different long-run predictions in learning models gained a widespread attention among experimental economists to understand how people play games by comparing their performance in experimental studies (Mookherjee and Sopher (1994), Erev and Roth (1998), Feltovich (2000), Altavilla et al. (2006)). The experimental studies find that although, under certain settings, aspiration-based learning models are superior to belief-based ones in terms of performance, in general the performance of these models depends on the games played and the environment in which such experiments are conducted.

The findings in the experimental studies motivated researchers to study the analytical properties of learning models and identify the possible explanations of differences observed in their performance. Sarin and Vahid (1999) show that learning, in a simple dynamic model of choice, which shares some features with both of these competing learning models, leads to maxmin choices. Hopkins (2002) finds the speed of learning as the main difference among these models. Dziubiński and Roy (2012) study the survival of the learning models in evolutionary settings and find conditions under which the exclusive use of either of these models are supported in the long-run.

In this paper, I take a different route, and instead of studying the predictive power of these learning models, I concentrate on the assumptions on which these models are built. Recent studies, in aspiration-based models, find that in the repeated prisoner's dilemma played by a large population of players, cooperation, a non-Nash outcome, is sustained in the long-run when players' aspiration level, in each round, is equal to a current average payoff in the population (Palomino and Vega-Redondo (1999), Dixon (2000), Oechssler (2002)). The results contradict the predictions of conventional belief-based models.

Although these contradictory results are often associated with different rationality assumptions present in the models (see Bendor et al. (2001)), in this paper I argue that different assumption with regard to information structures in the models are nevertheless important: in belief-based models, a player knows both her and her opponents' action spaces as well as her own payoff function, and observes her opponents' play in each round; in aspiration-based models, on the other hand, a payoff matrix of a game is unknown and each player observes only her own realized payoffs and the average payoff, but not her opponents' moves.

The aim of this paper is to show that, even though the rationality assumption of players differs in these models, belief-based models deliver the same predictions as aspiration-based models about the long-run outcome of the learning process, if studied under the informational settings similar to those in aspiration-based models. To this purpose, I develop a belief-based learning model of the prisoner's dilemma in an anonymous random-matching setting when the game is played repeatedly by a continuum of players, as well as studing its long-run properties.

To introduce informational assumptions similar to those in aspiration-based learning, and different from conventional belief-based models, I assume that players do not know the payoff matrix of the stage game and know only their own action set. In addition, players do not observe actions of other players, and hold beliefs not over the likely play of others, but over the payoffs they expect from different actions.[1] At the end of each round, payoffs are realized and each player observes her own payoff as well as the current average payoff in the population. Players perceive the observables as noisy private and public signals and update their beliefs about expected payoffs using Bayes' rule.

A general theory to study a class of single player infinitly repeated games with unknown payoff distributions is developed in Easley and Kiefer (1988). The main question in the models studied in Easley and Kiefer (1988) is whether players can learn true parameters in the limit. Easley and Kiefer (1988) show that though the player's beliefs about unknowns eventually converge due to learning, beliefs need not converge to truth. The reason for this observation is quite simple. In the models, covered in Easley and Kiefer (1988), information is generated endogenously and depends on the action choices of the players. Therefore, along any sample path for which beliefs do converge the sequence of actions may be converging as well and actions may not generate enough information to identify the true parameters. This may lead to incomplete learning as was found in previous studies, see, e.g., Rothschild (1974), McLennan (1984), Rustichini and Wolinsky (1995), Brezzi and Lai (2000).

In this paper, I extend the theory developed in Easley and Kiefer (1988) to the infinitly repeated games of the continuum of players with anonymous random matching. In contrast to the models studied in Easley and Kiefer (1988) where the main question is to identify conditions which guarantee convergence of the player's beliefs to "truth", here in this paper, I focus primarily on studying the long-run equilibrium outcomes of the learning process.

The results of the paper can be summarized as follows. First, I extend the theory of Easley and Kiefer (1988) to the continuum of players settings and define the equilibrium of the game. Second, to analyse the long-run equilibrium outcomes of the learning process I introduce the stochastic stability analysis. Finally, I apply the methods developed to analyse the learning process in the infinitly repeated prisoner's dilemma and study its long-run equilibrium outcomes.

To compare long-run predictions of the belief-based model in my paper with aspiration-based learning models, I study the limit behavior of the learning process when the precision of

[1] In Sarin and Vahid (2001) players also hold beliefs over payoff, but beliefs are scalar-valued and are updated using some deterministic rule, whereas beliefs in my model are vector valued and are updated using Bayes' rule.

signals improves significantly in the limit. When the public signals become perfectly precise in the limit, the stochastic stability analysis reveals that cooperation is the only stochastically stable state of the belief-based learning process which coincides with the predictions of aspiration based learning. The result highlights the importance of informational assumptions in decision problems under uncertainty and illustrates their essential role in determining the long-run outcome of the learning process.

The rest of the paper is organized as follows. The next section describes the model setup and players' decision problem. Section 3.3 studies the limit beliefs and polices and characterizes the long-run equilibrium outcome of learning dynamics. Section 3.4 concludes the paper.

3.2. The model

In each time period $t = 1, 2, \dots,$ a continuum of players with the measure one indexed by $i \in [0,1]$ are randomly and anonymously matched in pairs and play the prisoner's dilemma game, called the stage game with action set $A = \{C, D\}$ and payoff matrix as follows:

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

where C stands for "cooperation" and D for "defection". Let m be a measure-preserving matching function, defined as $m: [0,1] \to [0,1]$, such that $m(m(i)) = i$ and $m(i) \neq i$. In each time period matching is drawn uniformly and independently across time.

Let a payoff profile $\theta \in \Theta \subset R^4$ describe payoffs of the stage game. Then, nature, at the beginning of the game, draws $\theta \in \Theta$, according to a fixed distribution, and sets the stage game $G(\theta)$. The payoff profile θ is never revealed to players and also, players do not know the payoff structure of $G(\theta)$. Each player only knows her own action set A.

The set of possible outcomes of $G(\theta)$ coincides with the set of action profiles A^2 and is given as $A^2 = \{(C,C), (C,D), (D,C), (D,D)\}$. The conventional payoff constraints of the prisoner's dilemma, $g, l > 0$ and $-1 < g - l < 1$, are imposed, so that the dominant strategy of the stage game $G(\theta)$ is D and playing (C,C) yields a higher payoff than a symmetric randomization over the other feasible outcomes of the stage game $G(\theta)$.

For any set X, let $\Delta(X)$ denote a set of probability measures over X. Define a payoff function of $G(\theta)$ as $r\colon A^2 \times \Theta \to R$ and a player's behavior strategy, a probability distribution over her actions, as $\tilde{\pi}\colon A \to \Delta(A)$. Then, each player i's expected payoff from action $a_i \in A_i$ if matched with player $m(i)$, is defined as

$$\tilde{r}(a_i, \tilde{\pi}_{m(i)}, \theta) = \sum_{a_{m(i)} \in A} r(a_i, a_{m(i)}, \theta)\tilde{\pi}_{m(i)}(a_{m(i)}). \tag{3.1}$$

Let $\pi_i \in \{0,1\}$ denote player i's choice variable where $\pi_i = 1$ corresponds to action D and $\pi_i = 0$ corresponds to action C.

3.2.1. "Structural" model

True data generating process. At the end of each time period t, given the outcome of the stage game $G(\theta)$ and a realization of the exogenous random variable ε_{it}, each player i receives the payoff y_{it}, such that

$$y_{it} = r(a_{it}, a_{m(i),t}, \theta) + \sigma_\varepsilon \varepsilon_{it}, \tag{3.2}$$

where $\varepsilon_{it} \sim \mathcal{N}(0,1)$ is independent and identically distributed across players, and serially uncorrelated over time; σ_ε is a standard deviation.

Each player i, in addition, receives a noisy public signal y_t about the current average payoff in the population defined as

$$y_t = \int y_t(i)di + \sigma_\eta \eta_t, \tag{3.3}$$

where $\eta_t \sim \mathcal{N}(0,1)$ is common noise due to the aggregation of information, independent of both $\int y_t(i)di$ and ε_{it} for each $i \in [0,1]$; σ_η is a standard deviation. Given the assumptions of the model, due to the strong law of large numbers we have $\int \varepsilon_t(i)di = 0$ and the average payoff $\int y_t(i)di$ is defined as

$$\int y_t(i)di = \int r(a_t(i), a_t(m(i), \theta)di. \tag{3.4}$$

Perceptions and belief rules. Define player i's expected payoff from playing C as $\alpha_{it} = \tilde{r}(C, \tilde{\pi}_{m(i),t}, \theta)$ and the expected payoff gain from playing D as $\beta_{it} = \tilde{r}(D, \tilde{\pi}_{m(i),t}, \theta) - \tilde{r}(C, \tilde{\pi}_{m(i),t}, \theta)$.

As player i does not know the payoff matrix of $G(\theta)$ and does not observe the behavior strategy choices of other players, she holds the set of assessments Γ over the possible values of $(\alpha_{it}, \beta_{it})$. Let $\gamma_i = [\alpha_i, \beta_i]^\top$ be a random variable with the support Γ. Then, each player i perceives the stage game payoff, given the choice variable π_{it}, to be defined as

$$v(\pi_{it}, \gamma_i) = \alpha_i + \beta_i \pi_{it}. \tag{3.5}$$

Let $v_{it} \equiv v(\pi_{it}, \gamma_i)$ and $x_{it} \equiv [1, \pi_{it}]^\top$, then we can rewrite (3.5) as[2]

$$v_{it} = x_{it}^\top \gamma_i, \tag{3.6}$$

and the observed y_t and y_{it}, the noisy public and private signals, are given as

$$\begin{aligned} y_t &= v_{it} + \sigma_\eta \eta_t, \\ y_{it} &= v_{it} + \sigma_\varepsilon \varepsilon_{it}. \end{aligned} \tag{3.7}$$

Define player i's beliefs over Γ as $\mu_{it} \in \Delta(\Gamma)$. Each player i is assumed to use Bayes' rule to update her prior beliefs μ_{it-1} at each time period after the observations of (π_{it}, y_{it}, y_t) become available, such that, the posterior belief μ_{it}, for all $t = 1, 2, \ldots$, is given as

$$\mu_{it} = \phi(\mu_{it-1}, \pi_{it}, y_{it}, y_t), \tag{3.8}$$

where the belief rule ϕ represents the Bayes' rule operator.

For simplicity, assume that each player i, at the beginning of the game, holds improper uniform priors μ_{i0} over the whole real line. This implies that the belief rule ϕ in (3.8) is a well behaved continuous function. Each time period t, by standard arguments in Bayesian econometrics given the normality assumptions, see, e.g., Zellner (1970), the posterior of μ_{it-1} according to the belief rule ϕ, if $\lambda \equiv 1/(1 + \sigma_\eta^2/\sigma_\varepsilon^2)$ and $\sigma_\xi \equiv \sigma_\varepsilon / \sqrt{1 + \lambda/(1-\lambda)}$, is given as

$$\mu_{it} = \mathcal{N}(\hat{\gamma}_{it}, \Sigma_{it}) \tag{3.9}$$

[2] Players are assumed to consider the environment to be stationary. In the long-run, each player's beliefs converge, as we will see in the next section, and players actions do as well. Therefore, the possible misspecification due to stationarity assumption disappears in the limit, see, e.g, Bray and Savin (1986).

with the mean

$$\hat{\gamma}_{it} = (X_{it}^{\top} X_{it})^{-1} X_{it}^{\top} \tilde{Y}_{it} \tag{3.10}$$

and covariance matrix

$$\Sigma_{it} = \sigma_{\xi}^2 (X_{it}^{\top} X_{it})^{-1}, \tag{3.11}$$

where $\tilde{Y}_{it} = \lambda Y_t + (1-\lambda) Y_{it}$, $X_{it}^{\top} = [X_{it-1}^{\top}, x_{t-1}^{\top}]$, $Y_t^{\top} = [Y_{t-1}^{\top}, y_{t-1}^{\top}]$ and $Y_{it}^{\top} = [Y_{it-1}^{\top}, y_{it-1}^{\top}]$ are the matrices of observations.

We can express (3.10) and (3.11) in the recursive representation corresponding to (3.8), see, e.g., Zellner (1971), as

$$\begin{aligned} \hat{\gamma}_{it} &= \hat{\gamma}_{it-1} + \kappa_{it}(\tilde{y}_{it-1} - x_{it-1}^{\top} \hat{\gamma}_{it-1}), \\ \Sigma_{it} &= \Sigma_{it-1} - \Sigma_{it-1} x_{it-1} (x_{it-1}^{\top} \Sigma_{it-1} x_{it-1} + \sigma_{\xi}^2)^{-1} x_{it-1}^{\top} \Sigma_{it-1}, \end{aligned} \tag{3.12}$$

where $\tilde{y}_{it-1} = \lambda y_{t-1} + (1-\lambda) y_{it-1}$, and $\kappa_{it} = \Sigma_{it-1}(x_{it-1}^{\top} \Sigma_{it-1} x_{it-1} + \sigma_{\xi}^2)^{-1} x_{it-1}$.

We see from (3.12) that the information structure (3.7) is equivalent to receiving the signal $\tilde{y}_{it}$, which is a weighted average of the public and private signals, defined as

$$\tilde{y}_{it} = v_{it} + \sigma_{\xi} \xi_{it}, \tag{3.13}$$

where $\xi_{it} = (\sqrt{\lambda}\eta_t + \sqrt{1-\lambda}\varepsilon_{it}) \sim \mathcal{N}(0,1)$.[3] This implies that the belief rule (3.8) is equivalent to

$$\mu_{it} = \phi(\mu_{it-1}, \pi_{it}, \tilde{y}_{it}) \tag{3.14}$$

Information sets. Let $\mathcal{H}_t = \Delta(\Gamma) \times \Pi_{s=1}^{t-1}(\{0,1\} \times \mathbb{R} \times \Delta(\Gamma))$. At the end of the time period t, each player i knows her decision choice π_{it}, signal $\tilde{y}_{it}$ and updated prior beliefs μ_{it} according to Bayes' rule (3.14), but do not observe other players' decision choices, beliefs or realized payoffs. This defines player i's history as $h_{it} = \{h_{it-1}, (\pi_{it-1}, \tilde{y}_{it-1}, \mu_{it-1})\} \in \mathcal{H}_t$ with $h_{i0} = \{\mu_{i0}\}$.

[3] We have $\xi_{it} = (\lambda\sigma_{\eta}\eta_t + (1-\lambda)\sigma_{\varepsilon}\varepsilon_{it})/\sigma_{\xi}$, but after some algebraic manipulations, it simplifies into $\xi_{it} = \sqrt{\lambda}\eta_t + \sqrt{1-\lambda}\varepsilon_{it}$.

3.2.2. Decision problem

A player's decision problem is to maximize an expected discounted payoff over an infinite horizon. Given beliefs μ_{it}, each player i's expected payoff, $u: \{0,1\} \times \Delta(\Gamma) \to \mathbb{R}$, is defined as

$$u(\pi_{it}, \mu_{it-1}) = \int_{\Gamma} \int_{\mathbb{R}} \left(v(\pi_{it}, \gamma_i) + \sigma_{\varepsilon} \varepsilon_{it} \right) dv d\mu_{it-1}, \tag{3.15}$$

where v is a probability measure over a standard normal random variable.

Define a policy function $\psi_t: \mathcal{H}_t \to \{0,1\}$, which, given a history $h_t \in \mathcal{H}_t$, specifies the date t choice variable $\pi_t = \psi_t(h_t)$. Let each player i have the same discount factor $\delta \in (0,1)$. Then, given the policy, a sequence $\psi_i = \{\psi_{it}\}_{t=1}^{\infty}$, player i's expected discounted payoff generated by the policy ψ_i is equal to

$$I_{\psi_i}(\mu_{i0}) = \mathrm{E}\left[\sum_{t=1}^{\infty} \delta^t u(\pi_{it}, \mu_{it-1}) \right]. \tag{3.16}$$

Let Ψ be a set of all possible polices. Player i has to choose the policy $\psi_i \in \Psi$ that guarantees the maximum expected discounted payoff $\sup_{\psi_i \in \Psi} I_{\psi_i}(\mu_{i0})$. Then, define a value function $V(\mu_{i0})$ as

$$V(\mu_{i0}) = \sup_{\psi_i \in \Psi} I_{\psi_i}(\mu_{i0}). \tag{3.17}$$

DEFINITION 5. *A policy $\psi^* \in \Psi$ is optimal if $I_{\psi^*}(\mu_0) = V(\mu_0)$.*

It is easy to verify that the standard assumptions needed for the existence of the value function and optimal policy in the infinite horizon dynamic programming problem, such that, the compactness of space of the choice variable, the boundedness and concavity of the expected payoffs, the continuity of Bayes' rule operator are satisfied by the assumptions of the model (see, e.g., Blackwell (1965), Blume et al. (1982), Easley and Kiefer (1988)).

THEOREM 1 (Blume et al. (1982, Theorem 1.1)). *(i) There is a unique continuous value function $V: \Delta(\Gamma) \to \mathbb{R}$.*

(ii) There is a stationary optimal policy $\Pi: \Delta(\Gamma) \to \{0,1\}$ for the infinite horizon problem (3.17) *such that*

$$\Pi(\mu) = \left\{ \pi \in \{0,1\} \,\middle|\, V(\mu) = \max_{\pi \in \{0,1\}} \left\{ u(\pi, \mu) + \delta \int_{\Delta(\Gamma)} \int_{\mathbb{R}} V(\phi(\mu, \pi, \tilde{y})) dv d\mu \right\} \right\}.$$

In the next section, we will study the limit behavior of the learning process and characterize the optimal limit policies.

3.3. Limit beliefs, policies and equilibrium

In this section, we will study the limit behavior of the learning process. A general approach to studying the limit behavior of the posterior processes in a single player setting is provided in Easley and Kiefer (1988). Here, I will recall some of their results, but whenever appropriate I will provide different proofs specific to the model in the paper. In addition, I extend the work of Easley and Kiefer (1988) in two ways. First, I extend their analysis to multiplayer settings and discuss the equilibrium of the model. Second, I introduce stochastic stability analysis which functions as a selection mechanism over the elements of the set of limit beliefs. Finally, I find the combination of private and public information that leads to cooperation in the long-run.

3.3.1. Limit beliefs and policies

As is well known, beliefs generated by the Bayes' rule operator are bounded martingales, and by the martingale convergence theorem the process of beliefs generated by belief rule ϕ in (3.14), as it represents the Bayes' rule operator, converge to some limit beliefs $\mu_\infty \in M^\infty \subset \Delta(\Gamma)$.[4]

Theorem 2 (Easley and Kiefer (1988, Theorem 4)). *There exists $\mu_\infty \in \Delta(\Gamma)$ such that $\mu_t \xrightarrow{a.s.} \mu_\infty$.*

The convergence in Theorem 2 does not mean the convergence to the "truth" in the sense of the consistency results of OLS, see, e.g., Easley and Kiefer (1988). In classical econometrics, regressors are assumed to be exogenous, which guarantees convergence to true parameters if the regression model is specified correctly. Here, the regressors are endogenous and when beliefs converge, the regressors may well be converging as well and may not generate enough variability in the regressors to identify the true parameters. This may lead to incomplete learning, as was found in previous studies, see, e.g., Rothschild (1974), McLennan (1984), Rustichini and Wolinsky (1995).

[4] In what follows, whenever it is not confusing, I will omit the subscript i.

To characterize the set of limit beliefs M^{∞}, we will, first consider the conditions under which beliefs converge to "truth" and see that this holds indeed under the similar conditions needed for the consistency of OLS estimates, that is when a player's choice variable π_t does not converge, such that $\mathrm{Var}[\pi_t] > 0$. Secondly, we will see that when π_t does converge, $\mathrm{Var}[\pi_t] \to 0$, then the set of limit beliefs M^{∞} may not be a singleton and I will use stochastic stability analysis to identify a stochastically stable limit belief μ_{∞} and consequently the long-run outcome of the learning process.

For a given limit belief $\mu_{\infty} \in M^{\infty}$, a limit policy $\pi_{\infty} = \lim_{t\to\infty} \pi_t$ should be consistent with it, such that $\pi_{\infty} \in \Pi(\mu_{\infty})$. In addition, given the limit policy $\pi_{\infty} \in \Pi^{\infty}$, the limit beliefs $\mu_{\infty} \in M^{\infty}$ should be consistent with the observations such that the orthogonality condition $\mathrm{E}[x_t(\tilde{y}_t - x_t^{\top}\hat{\gamma})] = 0$ holds. If, under the limit beliefs the orthogonality condition does not hold and the limit beliefs are not consistent with the observations, then eventually players will change their beliefs through the belief rule, but then such beliefs could not be limit beliefs.

Lemma 1. *Let γ_{∞} denote a limit of the sequence $\{\hat{\gamma}_t\}$ such that $\gamma_{\infty} = \lim_{t\to\infty} \hat{\gamma}_t$. Then, A belief $\mu_t \xrightarrow{a.s.} \mu_{\infty}$ if and only if the orthogonality condition $\mathrm{E}[x_t(\tilde{y}_t - x_t^{\top}\gamma_{\infty})] = 0$ holds.*

Proof. Applying the matrix inversion lemma to (3.12), see, e.g., Anderson and Moore (1979), yields

$$\hat{\gamma}_{s+1} = \Sigma_{s+1}(x_s\tilde{y}_s + \Sigma_s^{-1}\hat{\gamma}_s).$$

After some algebraic manipulations, it simplifies into

$$\Sigma_{t+1}^{-1}(\hat{\gamma}_{t+1} - \hat{\gamma}_t) = x_t(\tilde{y}_t - x_t^{\top}\hat{\gamma}_t).$$

Let $M_t = \left\|\Sigma_{t+1}^{-1}(\hat{\gamma}_{t+1} - \hat{\gamma}_t)\right\|$ and $g_t = x_t(\tilde{y}_t - x_t^{\top}\hat{\gamma}_t)$. Fix the sample path for which beliefs converge. Given the sample path, we have $\mathrm{E}_{\mu_t}(M_t) \xrightarrow{a.s.} 0$ and $\mathrm{E}_{\mu_t}(g_t) \xrightarrow{a.s.} 0$. Then, as for each $t > 0$ we have $\|g_t\| \le M_t$, by the dominated convergence theorem we can move the limit operation inside the expectation operator yielding $\lim_{t\to\infty} \mathrm{E}_{\mu_t}[g_t] = \mathrm{E}[x_t(\tilde{y}_t - x_t^{\top}\gamma_{\infty})] = 0$.

The second part of the proof relies on the same steps as in the first part but in the opposite direction. □

We are particularly interested in the limit behavior of the learning process and thus, it suffices to characterize, for a given limit belief $\mu_\infty \in M^\infty$, limit choice π_∞ which is consistent with μ_∞, such that $\pi_\infty \in \Pi(\mu_\infty)$. Suppose that γ^* is the true parameter vector. Then, the true data generating process for $\tilde{y}_t$ is given as $\tilde{y}_t = x_t^\top \gamma^* + \varepsilon_t$.

LEMMA 2. *Let $\mathbf{1}_{\{\beta_\infty \geq 0\}}$ be an indicator function of playing D, equal 1 if $\beta_\infty \equiv \mathrm{E}_{\mu_\infty}[\beta_t] \geq 0$ and 0 otherwise. Then, limit choice π_∞, given limit beliefs μ_∞, is $\pi_\infty = \mathbf{1}_{\{\beta_\infty \geq 0\}}$.*

Proof. Fix the sample path for which beliefs converge. Given the sample path, Lemma 1 implies the orthogonality condition $\mathrm{E}[x_t(\tilde{y}_t - x_\infty{}^\top \gamma_\infty)] = 0$. This, as $\mathrm{E}[\tilde{y}_t - x_t^\top \gamma_\infty] = 0$, yields $\tilde{y}_\infty \equiv \mathrm{E}[x_t^\top \gamma^*] = x_\infty{}^\top \gamma_\infty$ and thus, $\tilde{y}_t = \tilde{y}_\infty + \varepsilon_t$. As $\tilde{y}_t$ is a white noise, observing $(\pi_\infty, \tilde{y}_t)$ provides no information and thus, implies $\phi(\mu_\infty, x_\infty, \tilde{y}_t) = \mu_\infty$.

Then, applying the results of Easley and Kiefer (1988, Lemma 4), we have that limit choice π_∞, given limit beliefs μ_∞, solves the one-period problem

$$\pi_\infty \in \underset{\pi \in [0,1]}{\arg\max}\ u(\pi, \mu_\infty).$$

Solving the one-period problem yields $\pi_\infty = \mathbf{1}_{\{\beta_\infty \geq 0\}}$. □

Next, we are interested in the conditions that guarantee the convergence of beliefs to the "truth" and also to characterize the set of limit beliefs when such conditions are not met. Let σ_π denote the limit of the sequence of standard deviations defined as $\sigma_\pi = \lim_{t \to \infty} \sqrt{\mathrm{Var}[\pi_t]}$. The following theorem determines the set of limit beliefs and the limit set of the sequence $\{\hat{\gamma}_t\}$.

THEOREM 3. *The sequence $\{\hat{\gamma}_t\}$ converges to the true parameter vector $\gamma_\infty = \gamma^*$ if the sequence $\{\pi_t\}$ does not converge $\sigma_\pi > 0$. The convergence of beliefs to the "truth" implies $\gamma_\infty = \mathrm{E}[x_t x_t^\top]^{-1} \mathrm{E}[x_t^\top y_t]$. Otherwise, if the sequence $\{\pi_t\}$ does converge $\sigma_\pi = 0$, the set of limit beliefs, given by $\Gamma^\infty = \left\{\gamma \in \Gamma \,\middle|\, x_\infty{}^\top (\gamma^* - \gamma) = 0\right\}$, might not be a singleton.*

Proof. Fix the sample path for which beliefs converge. Given the sample path, Lemma 1 implies the orthogonality condition $\mathrm{E}[x_t(\tilde{y}_t - x_t^\top \gamma_\infty)] = 0$. Then, as the true data generating process for $\tilde{y}_t$ is $\tilde{y}_t = x_t^\top \gamma^* + \varepsilon_t$, this yields $\mathrm{E}[x_t x_t^\top (\gamma^* - \gamma_\infty)] = 0$.

Consider first the case when $\sigma_\pi > 0$. This implies that $\mathrm{E}[x_t x_t^\top]$ is an invertible matrix. Then, the orthogonality condition yields a unique solution $\gamma_\infty = \gamma^* = \mathrm{E}[x_t x_t^\top]^{-1} \mathrm{E}[x_t^\top y_t]$.

Next, consider the case $\sigma_\pi = 0$. As $\mathrm{E}[x_t x_t^\top]$ is not invertible, the orthogonality condition has multiple solutions given by the set $\Gamma^\infty \equiv \{\gamma \in \Gamma \mid x_\infty^\top(\gamma^* - \gamma) = 0\}$. □

Theorem 3 implies that, though eventually beliefs converge due to learning, see Theorem 2, they may be "wrong". Hence, the orthogonality condition does not imply the consistency in the sense of OLS when regressors are endogenous and there is a need for additional analysis to identify the limit distribution of learning process.

3.3.2. Equilibrium

The fact that beliefs do converge allows us to talk about the meaningful equilibrium of the game, such as the self-confirming equilibrium, see, e.g., Fudenberg and Levine (1993), Cho and Sargent (2008). Using the results of Theorem 2, Lemma 1 and Lemma 2, we can define the equilibrium of the game as follows:

Definition 6. *A profile of a pair of limit choices and beliefs* $\prod_{i\in[0,1]} (\pi_{i\infty}, \mu_{i\infty}) \in \prod_{i\in[0,1]} \{0,1\} \times \Delta(\Gamma)$ *is the equilibrium of the game if for each* $i \in [0,1]$

(i) beliefs $\mu_{i\infty}$ *are consistent with the observations such that the orthogonality condition holds*

$$\mathrm{E}[x_{it}(\tilde{y}_{it} - x_{it}^\top \gamma_{i\infty})] = 0,$$

(ii) and given beliefs $\mu_{i\infty}$*, choice variable* $\pi_{i\infty}$ *solves the one-period problem such that*

$$\pi_{i\infty} \in \operatorname*{arg\,max}_{\pi\in\{0,1\}} u(\pi, \mu_{i\infty}).$$

As each player i's choice variable π_{it} is endogenous, when beliefs are converging, the choice variables might be converging as well, $\sigma_\pi = 0$, leading to the multiple solutions of the orthogonality condition (see Theorem 3) and thus, the equilibrium of the game should not be unique as well. To analyze the multiple equilibria, I consider stochastic stability analysis and study the long-run equilibrium of the game for $\sigma_\pi = 0$.

Stochastic stability and long-run dynamics. The system of belief rules $\tilde{\phi} = (\phi, \ldots, \phi)$ defines a Markov process $\mathcal{M}$ with a states space $\tilde{\Gamma} = \prod_{i\in[0,1]} \Gamma$. To characterize the long-run outcome of the learning process, we will study the support of the limit distribution of $\mathcal{M}$ for $\sigma_\pi \to 0$. The elements of the support of the limit distribution are called stochastically stable states.

Following Freidlin and Wentzell (1984), every stochastically stable state must be a limit state of the unperturbed dynamics and thus, it suffices to study the limit distribution of the reduced Markov process with states on the set of limit states of the unperturbed dynamics.[5]

A state $\tilde{\gamma} \in \tilde{\Gamma}$ is a limit state of the unperturbed dynamics, if a belief profile $\tilde{\mu}$ are invariant with respect to the system of belief rules $\tilde{\phi}$ given the state $\tilde{\gamma}$ for $\sigma_\pi = 0$. Theorem 3, for each player i, implies that the set of invariant beliefs is given by the orthogonality condition which defines the set of limit states $\Gamma^\infty = \left\{ \gamma \in \Gamma \,\middle|\, x_\infty^\top(\gamma^* - \gamma) = 0 \right\}$. Then, given the set of limit states Γ^∞, Lemma 2 implies that a player's choice variable π takes only two values, 1 and 0. This implies that each time period, we observe one of the following states: state CC in which players play C in all matches, state DD in which players play D in all matches and state CD in which some players play C and the others play D. Then, the reduced Markov process $\tilde{\mathcal{M}}$ is defined on a states space $\{CC, CD, DD\}$.

Proposition 9. *The support of the limiting distribution for $\sigma_\pi = 0$ contains only state CC if $\lambda = 1$ and otherwise, when $\lambda = 0$ it contains only state DD.*

Proof. See Appendix. □

The results of Proposition 9 show that the long-run equilibrium outcome depends on the weight, players assign to public and private signals, determined by the relative precisions of signals. When the public signals are perfectly precise, each player puts all weights on public signals $\lambda = 1$ in her signal extraction problem and the long-run outcome of the learning process is cooperation. However, when the private signals are perfectly precise, each player puts all weights on private signals $\lambda = 0$ in her signal extraction problem, and the long-run outcome of learning process is defection.

The predictions of aspiration-based learning models that the long-run outcome of the learning process is cooperation are obtained under the assumption of perfect public signals. The results thus show that the predictions of belief-based learning models coincide with the predictions of aspiration-based learning models when considered under the same informational assumptions as in aspiration-based learning models.

[5] The ideas of Freidlin and Wentzell are widely applied in game theory and further developed in Young (1993) and Kandori et al. (1993). Additional references and discussions are provided in Fudenberg and Levine (1998), and Vega-Redondo (2003).

3.4. Conclusion

This paper studies the long run outcomes of the belief-based learning process in the infinitely repeated prisoner's dilemma with anonymous random matching and unknown payoff distributions played by a continuum of players. The standard belief-based learning models predict the Nash equilibrium as the only long-run outcome of the learning process in games with a unique strict Nash equilibrium, such as a prisoners' dilemma. On the other hand, aspiration-based learning models allow dominated strategies to be played in the long run. The contradictory predictions of the learning models are often associated with a different level of rationality adopted in the models. However, in this paper, I show that an important role is, nevertheless, played by the informational assumptions. I find that the predictions of the belief-based learning models coincide with the predictions of aspiration-based learning as long as the public signals are perfectly precise and each player puts all weights on those signals. As a result, the only long-run outcome of the learning process is cooperation.

The results show that the predictions of the belief-based learning models coincide with the predictions of aspiration-based learning when the public signals are perfectly precise and each player puts all weights on public signals in her signal extraction problem and cooperation is the only long-run outcome of the learning process.

Appendix.

Proof of Proposition 9. Dropping i and time subscript t, the belief updating rule in (3.12) simplifies into

$$\hat{\gamma}' = \hat{\gamma} + \kappa(\tilde{y} - x^{\top}\hat{\gamma}). \tag{18}$$

Let $k_\infty = \lim_{t\to\infty} k_t$. If a' and a denote the actions played in the past and in the current period, respectively, then it is easy to verify that taking the limit $t \to \infty$ yields

$$k_\infty = \begin{cases} (0,0)^{\top} & \text{, if } a' = a \\ (0,1)^{\top} & \text{, if } a' = C \text{ and } a = D \\ (1,-1)^{\top} & \text{, if } a' = D \text{ and } a = C \end{cases} . \tag{19}$$

Then, given (18) and (19), players' beliefs are defined by

$$\hat{\gamma}' = \begin{cases} \hat{\gamma} & \text{, if } a' = a \\ (\hat{\alpha}, \tilde{y} - \hat{\alpha})^{\top} & \text{, if } a' = C \text{ and } a = D \\ (\tilde{y}, \hat{\alpha} + \hat{\beta} - \tilde{y})^{\top} & \text{, if } a' = D \text{ and } a = C \end{cases} . \tag{20}$$

As the player's choice variable π, given the limit states, takes only two values 1 and 0, for small but positive, σ_π there is a small but positive probability of switching $\tau \in [0,1]$ such that

$$\pi' = \begin{cases} \pi & \text{, with probability } 1 - \tau \\ 1 - \pi & \text{, with probability } \tau \end{cases} , \tag{21}$$

where π' is the player's choice next period. Given the switching probability τ, σ_π is defined as $\sigma_\pi = \sqrt{\tau(1-\tau)}$ and $\sigma_\pi = 0$ is equivalent to $\tau \to 0$.

Let $\tilde{\tau}$, defined as $\tilde{\tau} = \prod_{i \in [0,1]} \tau$, be a probability that in all of the matches players switch to an alternative action. Let $\hat{\tau}$, defined as $\hat{\tau} = \prod_{i \in [0,1]} (1 - \tau)$, be a probability that in all of the matches each player keeps playing the action she played in the past. Then, given (20), we can define a transition matrix $\mathbf{Q} = [q_{CC}, q_{CD}, q_{DD}]$ of the reduced Markov process $\tilde{\mathscr{M}}$ where

$$q_{CC} = \begin{cases} (1, 0, 0) & \text{, if } \lambda = 1 \\ (1 - 2\sqrt{\hat{\tau}\tilde{\tau}}, 2\sqrt{\hat{\tau}\tilde{\tau}}, 0) & \text{, if } \lambda = 0 \end{cases} , \quad q_{DD} = \begin{cases} (\tilde{\tau}, 1 - (\hat{\tau} + \tilde{\tau}), \hat{\tau}) & \text{, if } \lambda = 1 \\ (\tilde{\tau}, 0, 1 - \tilde{\tau}) & \text{, if } \lambda = 0 \end{cases} ,$$

$$q_{CD} = (0, 1 - 2\sqrt{\hat{\tau}\tilde{\tau}}, 2\sqrt{\hat{\tau}\tilde{\tau}}).$$

Then, when $\lambda = 1$, the state CC is an absorbing state of $\tilde{\mathscr{M}}$ and as it can be reached from any state with a non-zero probability, the support of limit distribution contains only state CC. Hence, CC is the only stochastically stable state if $\lambda = 1$.

When $\lambda = 0$, state DD is the only stochastically stable state as it can be reached from any state with a non-zero probability and a probability of leaving state DD is less than the same probability for other states.

□

References

Altavilla, C., L. Luini, and P. Sbriglia, "Social learning in market games," *Journal of Economic Behavior & Organization*, December 2006, *61* (4), 632–652.

Anderson, B.D.O. and J.B. Moore, *Optimal Filtering*, Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

Aumann, Robert and Adam Brandenburger, “Epistemic Conditions for Nash Equilibrium,” *Econometrica*, September 1995, *63* (5), 1161–1180.

Bendor, Jonathan, Dilip Mookherjee, and Debraj Ray, “Aspiration-Based Reinforcement Learning in Repeated Interaction Games: An Overview,” *International Game Theory Review*, 2001, *03* (02n03), 159–174.

Blackwell, David, “Discounted Dynamic Programming,” *The Annals of Mathematical Statistics*, 1965, *36* (1), 226–235.

Blume, Lawrence, David Easley, and Maureen O’Hara, “Characterization of optimal plans for stochastic dynamic programs,” *Journal of Economic Theory*, 1982, *28* (2), 221 – 234.

Bray, Margaret M and Nathan E Savin, “Rational Expectations Equilibria, Learning, and Model Specification,” *Econometrica*, September 1986, *54* (5), 1129–60.

Brezzi, Monica and Tze Leung Lai, “Incomplete Learning from Endogenous Data in Dynamic Allocation,” *Econometrica*, 2000, *68* (6), 1511–1516.

Cho, In-Koo and Akihiko Matsui, “Learning aspiration in repeated games,” *Journal of Economic Theory*, 2005, *124* (2), 171–201.

_ **and Thomas J. Sargent**, “Self-confirming equilibria,” in Steven N. Durlauf and Lawrence E. Blume, eds., *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan, 2008.

Dixon, Huw David, “Keeping up with the Joneses: Competition and the evolution of collusion,” *Journal of Economic Behavior & Organization*, 2000, *43* (2), 223–238.

Dziubiński, Marcin and Jaideep Roy, “Popularity of reinforcement-based and belief-based learning models: An evolutionary approach,” *Journal of Economic Dynamics and Control*, 2012, *36* (3), 433–454.

Easley, David and Nicholas M. Kiefer, “Controlling a Stochastic Process with Unknown Parameters,” *Econometrica*, 1988, *56* (5), 1045–1064.

Erev, Ido and Alvin E. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *The American Economic Review*, 1998, *88* (4), 848–881.

Feltovich, Nick, "Reinforcement-Based vs. Belief-Based Learning Models in Experimental Asymmetric-Information Games," *Econometrica*, 2000, *68* (3), 605–641.

Freidlin, M.I. and A.D. Wentzell, *Random Perturbations of Dynamical Systems*, New York: Springer-Verlag, 1984.

Fudenberg, D. and D. K. Levine, "Self-Confirming Equilibrium," *Econometrica*, 1993, *61* (3), 523–545.

Fudenberg, Drew and David K. Levine, *The Theory of Learning in Games*, MIT press, 1998.

_ and _ , "Learning and Equilibrium," *Annual Review of Economics*, 2009, *1* (1), 385–420.

Hopkins, Ed, "Two Competing Models of How People Learn in Games," *Econometrica*, 2002, *70* (6), 2141–2166.

Kandori, Michihiro, George J Mailath, and Rafael Rob, "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, January 1993, *61* (1), 29–56.

Karandikar, Rajeeva, Dilip Mookherjee, Debraj Ray, and Fernando Vega-Redondo, "Evolving Aspirations and Cooperation," *Journal of Economic Theory*, 1998, *80* (2), 292–331.

Kim, Youngse, "Satisficing and optimality in 2x2 common interest games," *Economic Theory*, 1999, *13*, 365–375.

McLennan, Andrew, "Price dispersion and incomplete learning in the long run," *Journal of Economic Dynamics and Control*, 1984, *7* (3), 331 – 347.

Mookherjee, Dilip and Barry Sopher, "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 1994, *7* (1), 62–91.

Oechssler, Jörg, "Cooperation as a result of learning with aspiration levels," *Journal of Economic Behavior & Organization*, 2002, *49* (3), 405–409.

Palomino, Frédéric and Fernando Vega-Redondo, "Convergence of aspirations and (partial) cooperation in the prisoner's dilemma," *International Journal of Game Theory*, 1999, *28* (4), 465–488.

Rothschild, M., "A Two-armed Bandit Theory of Market pricing," *Journal of Economic Theory*, 1974, *9*, 185–202.

Rustichini, A. and A. Wolinsky, "Learning about Variable Demand in the Long Run," *Journal of Economic Dynamics and Control*, 1995, *19*, 1283–1292.

Sarin, Rajiv and Farshid Vahid, "Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice," *Games and Economic Behavior*, 1999, *28* (2), 294 – 309.

__ **and** __ , "Predicting How People Play Games: A Simple Dynamic Model of Choice," *Games and Economic Behavior*, 2001, *34* (1), 104–122.

Vega-Redondo, Fernando, *Economics and the Theory of Games* Cambridge Books, Cambridge University Press, December 2003.

Young, H Peyton, "The Evolution of Conventions," *Econometrica*, January 1993, *61* (1), 57–84.

Young, H. Peyton, *Strategic Learning and its Limits*, Oxford, U.K.: Oxford University Press, March 2004.

Zellner, A., "Estimation of Regression Relationships Containing Unobservable Independent Variables," *International Economic Review*, October 1970, *11* (3), 441–54.

__ , *An introduction to Bayesian inference in econometrics* Wiley series in probability and mathematical statistics: Applied probability and statistics, J. Wiley, 1971.