

Yazar, Julide

**Article**

## Bayesian fictitious play in oligopoly: The case of risk-averse agents

Games

**Provided in Cooperation with:**

MDPI – Multidisciplinary Digital Publishing Institute, Basel

*Suggested Citation:* Yazar, Julide (2024) : Bayesian fictitious play in oligopoly: The case of risk-averse agents, Games, ISSN 2073-4336, MDPI, Basel, Vol. 15, Iss. 6, pp. 1-15, <https://doi.org/10.3390/g15060040>

This Version is available at:

<https://hdl.handle.net/10419/330109>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

Article

# Bayesian Fictitious Play in Oligopoly: The Case of Risk-Averse Agents

Julide Yazar 

Department of Economics, Ohio Wesleyan University, Delaware, OH 43015, USA; jjyazar@owu.edu

**Abstract:** A number of learning models have been suggested to analyze the repeated interaction of boundedly rational agents competing in oligopolistic markets. The agents form a model of the environment that they are competing in, which includes the market demand and price formation process, as well as their expectations of their rivals' actions. The agents update their model based on the observed output and price realizations and then choose their next period output levels according to an optimization criterion. In previous works, the global dynamics of price movement have been analyzed when risk-neutral agents maximize their expected rewards at each round. However, in many practical settings, agents may be concerned with the risk or uncertainty in their reward stream, in addition to the expected value of the future rewards. Learning in oligopoly models for the case of risk-averse agents has received much less attention. In this paper, we present a novel learning model that extends fictitious play learning to continuous strategy spaces where agents combine their prior beliefs with market price realizations in previous periods to learn the mean and the variance of the aggregate supply function of the rival firms in a Bayesian framework. Next, each firm maximizes a linear combination of the expected value of the profit and a penalty term for the variance of the returns. Specifically, each agent assumes that the aggregate supply of the remaining agents is sampled from a parametric distribution employing a normal-inverse gamma prior. We prove the convergence of the proposed dynamics and present simulation results to compare the proposed learning rule to the traditional best response dynamics.



**Citation:** Yazar, J. Bayesian Fictitious Play in Oligopoly: The Case of Risk-Averse Agents. *Games* **2022**, *15*, 40. <https://doi.org/10.3390/g15060040>

Academic Editors: Ulrich Berge, Kjell Hausken, Riccardo D. Saulle, Marco A. Marini and Giorgos Stamatopoulos

Received: 15 October 2024  
Revised: 13 November 2024  
Accepted: 24 November 2024  
Published: 27 November 2024



**Copyright:** © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** Bayesian learning; Cournot model; risk-averse learning; fictitious play; normal-inverse gamma prior

## 1. Introduction

In the dynamic Cournot quantity competition model, agents choose their output levels in each period based on the observed history of realized market prices up to the current period. In this stylized model, there exists a single good, and firms make their supply decisions simultaneously at each period before the market price is realized, as determined by the total quantity of output supplied by the firms in the market. It is widely recognized that the Cournot–Nash equilibrium is the central solution concept in oligopoly theory. In Cournot–Nash equilibrium, no firm has an incentive to unilaterally deviate from its equilibrium strategy, as such deviations are unprofitable. However, the adjustment process in which the firms in an oligopoly reach the Cournot–Nash equilibrium is not always straightforward. In practice, firms are complex organizations that rely on specific behavioral and adaptation rules to make decisions about their production quantities and pricing, while making predictions about the future actions of their rivals. To gain insights into the adaptation process that leads to Cournot–Nash equilibrium, previous work in this area has hypothesized different mechanisms for individual learning, where agents learn to play by introspection. Firms build an individualized model of the market, form expectations of rival firms' actions, and choose the best action given their assumptions. A prime example of individual learning is best response dynamics, where agents know the demand function

and compute the best output level to maximize their profits, assuming that the rival firms will repeat their actions from the previous period [1]. For a duopoly, best response learning converges to the Cournot–Nash equilibrium. However, Theocharis [2] showed that best response dynamics do not converge in oligopolies with three or more firms.

Contrary to the predictions of best response dynamics theory, experimental data from stylized laboratory experiments demonstrate convergent behavior consistent with the Cournot–Nash equilibrium, even in situations where theoretical models predict that the best response dynamics should theoretically diverge [3–5]. This suggests that, even when the NE is theoretically unstable under best response dynamics, participants in experimental settings may still converge to it. This behavior could be due to participants following different learning and adaptation rules or other behavioral factors not captured by best response dynamics. These findings have stimulated research into alternative learning models for oligopolies [6–8].

In all prior models of individual learning, firms act myopically, ignoring the *uncertainty* in their estimates of their rivals' actions and demand parameters. This approach can be partly justified by observing that risk-neutral agents choose their actions to maximize the expected profit. With a linear demand function, the expected profit depends only on the mean of the rival's supply estimate and not on its variance. However, this rationale does not apply to risk-averse agents, who seek to balance maximizing the expected profits with minimizing the variance in their returns. In this paper, we address this gap in the literature. We study learning by *risk-averse* firms in oligopolies, the resulting dynamics, and the convergence to market equilibrium. In our model, risk-averse players form and update Bayesian models for their rivals to estimate *both the mean and the variance* of their rivals' production decisions in a rigorous, principled framework, and they make output decisions based on these predictions.

Oligopoly theory often assumes that firms are risk-neutral, yet this assumption has limited validity as it does not always align with real-world behavior. Various factors drive firms to adopt risk-averse behaviors. For instance, under increased uncertainty, firms tend to reduce their capital investments to avoid potential losses, as shown by empirical studies linking greater uncertainty with lower investment levels [9]. Additionally, firms with non-diversified owners may display risk aversion to protect their concentrated holdings [10]. Risk-averse managers, particularly those with performance-based compensation, may also favor less risky strategies [11]. Finally, risk-averse firms frequently diversify and hedge to manage exposure, with evidence indicating that corporate hedging is a common method of mitigating risk [12]. Despite the empirical evidence of risk-averse behaviors in markets, the operational theory of risk-averse oligopoly has been relatively underdeveloped. Notably, in a single-period static setting, Asplund [13] explored how varying degrees of risk aversion and demand or cost uncertainty impact competition in oligopolistic markets. The analysis shows that firms' best response strategies under demand uncertainty decrease with higher risk aversion, softening the competition in the Cournot oligopoly model. However, these observations do not generalize to the dynamic setting, where firms compete repeatedly in the market.

*Related Work:* Our model is inspired by early work by Kirman [14–16]. In this work, risk-neutral firms combine their uncertainty regarding their rivals' output decisions into a single unknown parameter that encodes the market clearing price or the intercept of the linear market demand curve. When agents notice differences between their chosen output levels and the resulting prices, they update their linear demand model accordingly. The agents learn to play by forming and updating a Bayesian estimate of the unknown demand function that they face. In more recent work, Bischi et al. [17,18] extended this framework to nonlinear demand functions, using a local monopolistic approximation (LMA) for the demand. Firms believe that they face a linear market demand, locally approximating the actual nonlinear demand. In LMA, firms adjust their output based on experimental price elasticity estimates, trusting that these estimates are accurate. When their outcomes differ from their expectations, firms again revise their beliefs, triggering an adaptation process

that refines both their market understanding and strategies. This adaptation process can be shown to converge to the Cournot–Nash equilibrium. In general, the information that firms have about the market demand may be incomplete in many other respects: the functional form of the demand, the slope of the demand, or the market saturation point. For a broad discussion of the various learning models with misspecified and uncertain demand functions, we refer to the excellent survey in [8]. It is worth noting that learning in this setup focuses solely on estimating the mean output levels and not variances and therefore does not offer insights into the outcomes of dynamic interactions involving risk-averse agents.

Our work also extends classical fictitious play to the Bayesian setting. As an alternative individual learning model, classical fictitious play [19] assumes that each subject would take the empirical distribution of the actions chosen in past periods by each of his opponents as his belief about an opponent’s mixed strategy, and, in every period, he would play the best response to this belief when choosing his current strategy. Monderer and Shapley [20] show that, for “finite weighted potential games”—a class of games that includes the Cournot game—every fictitious play process converges in beliefs to equilibrium. However, the tabulation and memorization of all previous actions of rival firms is not a reasonable behavioral assumption for actual firm behavior. Recent theoretical work has shown that firms’ memory significantly affects the learning dynamics and might affect the convergence to the Cournot–Nash equilibrium [21–23]. Given the crucial role that memory plays in oligopolistic behavior, it is important to investigate individual learning algorithms that operate with limited memory and yet lead to convergence to the strategic Cournot–Nash equilibrium.

*Contributions:* In this paper, we propose an alternative learning model for risk-averse agents in which agents construct Bayesian models of their rivals. They update these models using rigorous conjugate priors and choose their actions according to the posterior predictive distribution. Each agent forms expectations of the mean and variance of profits in the next period based on the learned models of rival behavior. They then choose their production levels to maximize their profits, optimizing the trade-off between maximizing the expected profit and minimizing the variance in their returns. In this setup, firms naturally build heterogeneous models of rivals’ behavior despite observing the same market outcomes. This heterogeneity arises due to two factors. First, they have different priors upon entering the Cournot game. Second, each firm models the aggregate output of different sets of players. In this new setup with risk-averse agents performing Bayesian learning, we prove convergence to the Cournot–Nash equilibrium, even for cases where the best response dynamics are known to diverge. This extends the previous convergence results for individual learning given in [8,15,24] to  $M$  risk-averse agents, where the agents’ beliefs and actions follow a state evolution based on a *time-varying* state transition matrix.

In our proposed model, agents create a full *posterior distribution* for the rivals’ supply decisions in the next period. They use Bayes’ rule to update the parameters of the prior distribution, which is chosen to be a normal-inverse gamma distribution. The resulting posterior distribution allows us to model learning by risk-averse agents, who seek a trade-off between maximizing the expected profit and minimizing the variance in future returns. All prior work discussed above considers risk-neutral agents who maximize the expected value of their profits. To the best of our knowledge, this is the first work that integrates Bayesian risk estimates into the learning process for continuous-strategy oligopoly models. We note that the normal-inverse gamma distribution has been previously employed in modeling rewards in discrete-choice Markov games and Q-learning [25], but not in the context of fictitious play learning. Our Bayesian learning model provides a *memory-efficient* implementation of fictitious play that leads to convergence to the strategic Cournot–Nash equilibrium while keeping track of only a small set of parameters encoding the posterior distribution for the rivals’ decisions.

The rest of this paper is structured as follows. Section 2 introduces the Cournot model and reviews the best response and fictitious play dynamic adjustment processes as examples of agents’ formation of beliefs. In Section 3, we present the proposed learning

model based on Bayesian reasoning for the posterior estimation of a rival's supply decisions using a fixed known variance for the variations in output levels and provide proof of the local convergence of the resulting learning process to the fixed point of the iterations. In Section 4, we extend the analysis to the more general setup where agents track and estimate the variance of their rivals' actions and show convergence to the Cournot–Nash equilibrium. We illustrate the convergence of the global dynamics using simulations and provide comparisons with the best response dynamics. We conclude with directions for future research.

## 2. Cournot Model and Previous Work

We consider a market with a homogeneous good with  $M$  firms. The inverse demand function  $f(Q)$  determines the market price  $P_t$  at time  $t$  given the aggregate supply  $Q_t$  of the good from the  $M$  firms.

$$\begin{aligned} P_t &= f(Q_t) = a - bQ_t \\ &= a - b \sum_{i=1}^M q_{i,t} \end{aligned} \quad (1)$$

Each firm is assumed to have a monotonically increasing and concave total cost function  $C_i(q_{i,t})$ . Let  $q_t = [q_{1,t}, q_{2,t}, \dots, q_{M,t}]$  denote the vector of the output levels chosen by the  $M$  firms in period  $t$ . Then, the profit for each firm is given by

$$\pi_i(q_{i,t}, q_{-i,t}) = (f(Q_t)q_{i,t} - C(q_{i,t})) \quad (2)$$

Learning in the Cournot model can be expressed as each agent forming beliefs about the output  $q_{-i,t}$  that will be offered by the rival firms in period  $t + 1$ , based on the previous market history  $[q_{-i,1}, q_{-i,2}, \dots, q_{-i,t}]$ , and adjust their own output accordingly. In the following, we review two of these learning processes. Note that the realized market price depends only on the aggregate output level of the rival players  $Q_{-i,t} = \sum_{j \neq i} q_{j,t}$ . Therefore, players need to reason only about their rival's aggregate output.

### 2.1. Previous Learning Models and Risk-Neutral Agents

Previous works have considered the case of risk-neutral agents maximizing their expected profits after forming various learning models, which differ in how the agents form their estimates of the aggregate output level.

#### 2.1.1. Best Response Dynamics

Best response dynamics is the learning process suggested by Cournot in his seminal work on duopoly analysis. Under the best response dynamics, each firm assumes that the others will maintain their output level from the previous period and chooses its own output level as the best response to the supply of others in the previous period.

$$q_{i,t} = \max_{q_i} \pi(q_i, \hat{q}_{-i,t}) \quad (3)$$

$$= \max_{q_i} \pi(q_i, q_{-i,t-1}) \quad (4)$$

For the simple model of the quadratic cost function,

$$C_i(q_{i,t}) = c_i q_{i,t}^2 + d_i q_{i,t} \quad (5)$$

profit optimization results in the following dynamic update assuming an interior point solution ( $q_{i,t} > 0$ ):

$$q_{i,t} = \frac{a - d_i}{2(b + c_i)} - \frac{b}{2(b + c_i)} Q_{-i,t-1} \quad (6)$$

If we stack the quantity for each firm into a column vector  $q(t)$ , then the best response dynamics is given by the state space equation

$$q_t = C + Aq_{t-1}, \tag{7}$$

where  $C$  is a constant vector whose elements are given by  $\frac{a-d_i}{2(b+c_i)}$ , and  $A$  is an  $M \times M$  matrix with zeros on the diagonal and  $-\frac{b}{2(b+c_i)}$  for the off-diagonal terms. To study the dynamic stability of this state space difference equation, we can analyze the eigenvalues of the state transition matrix  $A$ . For example, for the simplest case of zero cost, the eigenvalues of the matrix  $A$  can be calculated easily as  $-(M - 1)/2$  and  $0.5$  (repeated  $M - 1$  times). Since the spectral radius is larger than 1 for  $M > 2$ , this dynamical system is not convergent except for duopoly. This instability result prompted modified theories of the best response dynamics where an inertia term was included so that the output level chosen by each firm is a weighted average between their previous output level choice and the best response to the aggregate output of the rivals:

$$q_t = \alpha q_{t-1} + (1 - \alpha)(C + Aq_{t-1}) \tag{8}$$

$$= (1 - \alpha)C + (\alpha I + (1 - \alpha)A)q_{t-1} \tag{9}$$

The resulting state transition matrix  $\alpha I + (1 - \alpha)A$  has a spectral radius of less than 1 for sufficiently large inertia  $\alpha$ . As a result, the difference equation converges to the Cournot equilibrium for this model. Empirical work and laboratory experiments provide some evidence for the inertia term in seller output behavior [26]; however, the best response dynamics suggest negative auto-correlation in the aggregate output levels, contradicting experimental data showing positive auto-correlation [27].

### 2.1.2. Simple Fictitious Play

Fictitious play is an alternative learning strategy, originally formulated for discrete action spaces, that postulates that each player tabulates the relative frequency of the past choices of its rivals and then chooses its own action to optimize the expected payoff, assuming that the rival will play a mixed strategy given by the relative frequency of past actions. In the case of the Cournot game, player  $i$  can tabulate the choices of its rivals using an empirical distribution:

$$q_{j,t} \sim \sum_{s=1}^{t-1} \frac{1}{t-1} \delta(q - q_{j,s}), \tag{10}$$

where  $\delta(q - q_0)$  is the Dirac delta function representing the unit mass at  $q_0$ . The expected value of the aggregate output by rival firms can then be calculated as

$$E[\hat{Q}_{-i,t}] = \sum_{j \neq i} \sum_{s=1}^{t-1} \frac{1}{t-1} q_{j,s} \tag{11}$$

Again, for the case of a linear demand function, the optimal output level can be derived as

$$q_{i,t} = \frac{a - d_i}{2(b + c_i)} - \frac{b}{2(b + c_i)} \sum_{j \neq i} \sum_{s=1}^{t-1} \frac{1}{(t-1)} q_{j,s} \tag{12}$$

For this learning model, the current state  $q_t$  depends on all of the previous output levels up to period  $t$ . Therefore, the state evolution can be only represented by an infinite set of difference equations [28]:

$$q_{t+1} = F\left(\sum_s A_{t,s} q_s\right), \tag{13}$$

where  $A$  is an infinite matrix with elements  $A_{t,s}$ , and the matrix satisfies two conditions: the elements in the upper diagonal  $A_{t,s} = 0$  for  $s > t$ , and each row sums to 1, i.e.,  $\sum_s A_{t,s} = 1$ .

For the case of fictitious play learning,  $A_{t,s} = 1/t$  for all  $s \leq t$ . This fictitious play dynamical model can be shown to be stable and converges to the Cournot–Nash equilibrium [28]. However, this model has two shortcomings. First, it does not account for the agents' prior beliefs at the beginning of the Cournot game. Second, the agents track only the moving average of their rivals' output choices, without estimating the expected future variability around this average—a crucial factor for risk-averse agents who are concerned about the variance of their profits. In the following, we propose an extended fictitious play model rooted in Bayesian reasoning that captures the heterogeneous prior beliefs of agents in the form of conjugate priors. In this framework, the output level of rival firms is modeled probabilistically, and both the mean and the variance of the posterior predictive distribution are used to formulate optimization criteria for risk-averse agents seeking a trade-off between maximizing the expected reward and minimizing the variance.

### 3. Bayesian Fictitious Play and Risk-Averse Agents

As before, we consider the oligopoly model with  $M$  firms, but the firms adopt a Bayesian approach in modeling and predicting their rivals' actions using a *continuous-valued* distribution. In particular, we assume that the deterministic demand function given in (1) is common knowledge and each firm  $i$  believes that the rival firms are using a stationary mixed (random) strategy where their aggregate output level  $Q_{-i}$  is sampled from a normal distribution  $Q_{-i,t} \sim \mathcal{N}(\mu_i, 1/\lambda)$  with an unknown mean  $\mu_i$  and known precision (inverse variance)  $\lambda$ :

$$\begin{aligned} P_t = f(Q_t) &= a - bq_{i,t} - bQ_{-i,t} \\ &= a - bq_{i,t} - b(\mu_i + \zeta_t) \end{aligned} \quad (14)$$

where random disturbances  $\zeta_t$  on the rival quantity choices are sampled from a normal distribution with zero mean and variance  $1/\lambda$ .

We assume that the players observe the realized market price but not directly the quantity choices of their rivals. In this setup, the firms cannot differentiate between the seemingly stochastic quantity choices of their rivals and potential random demand shocks that shift the inverse demand, as they only observe the aggregate effect on the market price. Therefore, an *alternative interpretation* of the fictitious play model is that firms model the aggregate choices of their rivals as an unknown constant  $\mu_i$ , observed through random demand shocks independently sampled from a normal distribution with *zero* mean and variance  $\sigma_t^2$ :

$$\begin{aligned} P_t = f(Q_t) &= a - bq_{i,t} - bQ_{-i,t} + \epsilon_t \\ &= a - bq_{i,t} - b\mu_i + \epsilon_t \\ &= a - bq_{i,t} - b\left(\mu_i + \frac{1}{b}\epsilon_t\right) \end{aligned} \quad (15)$$

The two interpretations in (14) and (15) are equivalent if  $\epsilon_t = b\zeta_t$  or  $\lambda = \frac{b^2}{\sigma_t^2}$ . We extend the analysis to the case of unknown precision  $\lambda$  in the next section.

In the following, for clarity, we will base our discussion on the mixed strategy play interpretation, where fictitious play learning corresponds to each firm updating their beliefs  $\mu_i$  from the observation of market price realizations using Bayesian reasoning. Firm  $i$  assumes that the distribution of the hypothesized aggregate output level  $\hat{Q}_{-i}$  of the remaining firms is given by

$$P(Q_{-i,t}|\mu_i) = \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{\lambda}{2}(Q_{-i} - \mu_i)^2\right) \quad (16)$$

The prior belief of firm  $i$  about the remaining firms' mean output level  $\mu_i$  is given by the conjugate normal distribution [29] prior:

$$P(\mu_i | v_{i,0}, \tau_{i,0}) = \sqrt{\frac{\tau_{i,0}}{2\pi}} \exp\left(-\frac{\tau_{i,0}}{2}(\mu_i - v_{i,0})^2\right) \tag{17}$$

The hyperparameters  $v_{i,0}, \tau_{i,0}$  are known and represent the prior beliefs of firm  $i$  about its rivals' behavior at the start of the Cournot game, at time 0. As the output choices are revealed in the game across time, Bayesian posterior estimation is used to update these hyperparameters, as detailed in [29], using

$$v_{i,t} = \frac{\tau_{i,t-1}v_{i,t-1} + \lambda Q_{-i,t}}{\tau_{i,t} + \lambda} \tag{18}$$

$$\tau_{i,t} = \tau_{i,t-1} + \lambda = \tau_{i,0} + t\lambda \tag{19}$$

Because of the choice of the conjugate prior, the posterior predictive distribution for  $\hat{Q}_{-i,t+1}$  (belief about the rivals' aggregate output level at time  $t + 1$ , based on all the output levels observed in the game up to time  $t$ ) is a Gaussian distribution with mean  $v_{i,t}$  and precision  $\frac{\tau_{i,t}\lambda}{\tau_{i,t} + \lambda}$ :

$$(\hat{Q}_{-i,t+1} | \{Q_{-i,s}\}_{s=1:t}) = \sqrt{\frac{\tau_{i,t}\lambda}{2\pi(\tau_{i,t} + \lambda)}} \exp\left(-\frac{\tau_{i,t}\lambda}{2(\tau_{i,t} + \lambda)}(Q_{-i,t} - v_{i,t})^2\right) \tag{20}$$

As a result, the expected value and variance of  $\hat{Q}_{-i}$  are given by

$$E[Q_{-i,t+1}] = v_{i,t} \quad V[Q_{-i,t+1}] = \frac{\tau_{i,t} + \lambda}{\tau_{i,t}\lambda} \tag{21}$$

Now, we introduce the second key component of our model: risk-averse agents who choose their output levels while trying to balance the trade-off between the expected value and the variance in their profit. The agents choose their output decisions to maximize a linear combination of the expected profit and a term that penalizes variance in the returns.

$$U(q_{i,t+1}) = E[\pi_i(q_{i,t+1}, q_{-i,t+1})] - \gamma V[\pi_i(q_{i,t+1}, q_{-i,t+1})] \tag{22}$$

$$= (a - b(q_{i,t+1} + E[Q_{-i,t+1}]))q_{i,t+1} - c_i(q_{i,t+1}) - \gamma b^2 q_{i,t+1}^2 V[Q_{-i,t+1}] \tag{23}$$

$$\tag{24}$$

Therefore, the optimal output choice for the risk-averse firm  $i$  is given by

$$q_{i,t+1} = \frac{a - d_i - bE[Q_{-i,t+1}]}{2(b + c_i + \gamma b^2 V[Q_{-i,t+1}])} \tag{25}$$

Inserting the Bayesian update equations for the expected value and variance of the profit from Equation (21), we obtain

$$q_{i,t+1} = \frac{a - d_i - bv_{i,t}}{2(b + c_i + \Gamma_{i,t})} \quad \text{with } \Gamma_{i,t} = \gamma b^2 \frac{\tau_{i,t} + \lambda}{\tau_{i,t}\lambda} \tag{26}$$

Combining this with the update equation of the second set of state variables,  $v_{i,t}$ , we have

$$v_{i,t} = \frac{(\tau_{i,0} + t\lambda)v_{i,t-1} + \lambda \sum_{j \neq i} q_{j,t}}{\tau_{i,0} + (t + 1)\lambda} \tag{27}$$

Recall that the parameters  $\tau_{i,0}$  and  $v_{i,0}$  encode the initial priors that each firm has about the behavior of its rivals. If we assume that the firms have no prior information about their

rivals at the start of the game, we can use an improper distribution with  $\tau_{i,0}$  approaching zero [29]. Then, the update equations take the simpler form

$$q_{i,t+1} = \frac{a - d_i - bv_{i,t}}{2(b + c_i + \Gamma_t)} \quad \text{with } \Gamma_t = \gamma b^2 \frac{t+1}{t\lambda} \tag{28}$$

$$v_{i,t} = \frac{(t)v_{i,t-1} + \sum_{j \neq i} q_{j,t}}{t+1} \tag{29}$$

The two sets of difference Equations (28) and (29) describe the behavior of the non-autonomous dynamical system. The system is non-autonomous, because the difference equations describing state evolution depend on  $t$ ; therefore, it is time-varying.

Comparing Equation (28) with the update equation for the classical fictitious play given in (12), we observe that, while the classical fictitious play uses a simple average of the previous output levels to estimate the current period’s output level, Bayesian fictitious play constructs its estimate using the weighted average of the prior beliefs and output data from the previous periods. Moreover, the optimal output levels are dampened by a time-evolving constant  $\Gamma_{i,t}$ , as a result of risk aversion. In the following, we show that the non-autonomous dynamic system that models Bayesian fictitious learning converges to a fixed point.

**Theorem 1.** Consider  $M$  firms with identical cost structures  $c_i = c$  and  $d_i = d$  for every firm  $i$ . Then, for the dynamical system governed by the state update equations given in (28) and (29), the fixed point given by

$$q_i = \frac{a - d}{2b + (M - 1)b + 2c + 2b^2 \frac{\gamma}{\lambda}} \quad v_i = \frac{(M - 1)(a - d)}{2b + (M - 1)b + 2c + 2b^2 \frac{\gamma}{\lambda}} \tag{30}$$

is locally asymptotically stable, if  $b/2(b + c + \gamma b^2/\lambda) < 1$ , i.e., if the initial beliefs are initialized in a small neighborhood of the fixed point, the fictitious play converges to the equilibrium given by the fixed point in the limit as  $t$  approaches infinity.

**Remark 1.** Note that the parameter  $c$  in the cost function can be negative; therefore, the condition  $b/2(b + c + \gamma b^2/\lambda) < 1$  is not always automatically satisfied.

**Remark 2.** With  $\gamma = 0$ , i.e., risk-neutral agents, the fixed point corresponds to the Cournot–Nash equilibrium. We observe that the increasing risk aversion of agents (with increasing  $\gamma$ ) improves the stability condition, as it becomes easier to satisfy  $b/2(b + c + \gamma b^2/\lambda) < 1$ .

**Proof.** To simplify the discussion and improve the exposition, we consider the case of two firms. We denote the weighted mean of the previous output decisions of firm 1 forming the belief of firm 2 as  $\bar{q}_{1,t} = v_{2,t}$  for clarity and in order to make the notation consistent with the prior work of [14–16,30]. Similarly, we denote the weighted mean of the previous output decisions of firm 2 forming the belief of firm 1 as  $\bar{q}_{2,t} = v_{1,t}$ . Then, the Bayesian fictitious play iterates given in (28) are simply given by

$$q_{1,t} = \frac{a - d - b\bar{q}_{2,t}}{2(b + c + \Gamma_t)} \quad q_{2,t} = \frac{a - d - b\bar{q}_{1,t}}{2(b + c + \Gamma_t)} \tag{31}$$

and we note that the weighted means  $(\bar{q}_{1,t}, \bar{q}_{2,t})$  of the previous output levels are given by

$$\bar{q}_{1,t} = \frac{t}{t+1} \bar{q}_{1,t-1} + \frac{1}{t+1} q_{1,t-1} \tag{32}$$

$$\bar{q}_{2,t} = \frac{1}{t+1} \bar{q}_{2,t-1} + \frac{t}{t+1} q_{2,t-1} \tag{33}$$

Inserting (31) into (32), we obtain

$$\bar{q}_{1,t} = \frac{t}{t+1} \bar{q}_{1,t-1} - \frac{1}{(t+1)} \frac{b}{2(b+c+\Gamma_t)} \bar{q}_{2,t-1} + \frac{1}{(t+1)} \frac{a-d}{2(b+c+\Gamma_t)} \tag{34}$$

$$\bar{q}_{2,t} = -\frac{1}{(t+1)} \frac{b}{2(b+c+\Gamma_t)} \bar{q}_{1,t-1} + \frac{t}{t+1} \bar{q}_{2,t-1} + \frac{1}{(t+1)} \frac{a-d}{2(b+c+\Gamma_t)} \tag{35}$$

We observe that the state evolution can be modeled as *linear* time-variant difference equations:

$$\bar{q}_{t+1} = C(t) + A(t)(\bar{q}_t) \tag{36}$$

The time-variant nature of the map  $A(t)$  and  $C(t)$  is a result of the direct dependence on  $t$ , as well as through  $\Gamma_t$ . We observe that the matrix  $A(t)$  can be diagonalized using eigenvalue decomposition:

$$\begin{aligned} A(t) &= \begin{bmatrix} \frac{t}{t+1} & -\frac{1}{(t+1)} \frac{b}{2(b+c+\Gamma_t)} \\ -\frac{1}{(t+1)} \frac{b}{2(b+c+\Gamma_t)} & \frac{t}{t+1} \end{bmatrix} \\ &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 - \frac{1+\rho_t}{t+1} & 0 \\ 0 & 1 - \frac{1-\rho_t}{t+1} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \end{aligned}$$

where  $\rho_t = \frac{b}{2(b+c+\Gamma_t)}$ . The convergence to the fixed point given by

$$q_1^* = q_2^* = \bar{q}^* = \frac{(a-d)}{3b+2c+2b^2\gamma/\lambda'} \tag{37}$$

is governed by the eigenvalues of the state transition matrix  $A(t)$  [31]. Specifically, the fixed point is (locally) asymptotically stable if

$$\prod_{t=t_0}^{\infty} \sqrt{\text{eig}_{\max}(A(t)^T A(t))} = \prod_{t=t_0}^{\infty} \max\{|\text{eig}(A(t))|\} = |1 - \frac{1-\rho_t}{t+1}|$$

for some  $t_0$ . This product converges to zero as long as  $\rho_t < 1$  for all  $t > t_0$  or when  $b < 2(b+c+\Gamma_t)$  for  $t$  larger than  $t_0$ . The inequality holds true for all  $t$ , if  $b < 2(b+c+\gamma b^2/\lambda)$  as  $\Gamma_t > \gamma b^2/\lambda$ , establishing the desired result. Observe that this inequality is not always satisfied, as the cost parameter  $c$  may be negative. Finally,  $b < 2(b+c)$  is a necessary condition for  $b < 2(b+c+\Gamma_t)$ . In other words, the condition on the eigenvalues ensures that the repeated application of  $A(t)$  is a contraction and therefore the fixed point of the iteration is asymptotically stable. The extension of the proof to a number of firms  $M > 2$  follows the same argument. The eigenvalues of the state transition matrix  $A(t)$  will have the eigenvalue  $1 - \frac{1+M\rho_t}{t+1}$  and  $(M-1)$  repeated eigenvalues of  $1 - \frac{1-\rho_t}{t+1}$ , with the latter being the larger of the two. Therefore, the same condition guarantees asymptotic local stability. □

Thus far, we have assumed that the variance (or, equivalently, the precision  $\lambda$ ) is known and fixed throughout the game. However, one may conjecture that, as the learning progresses, the output choices of rival firms will stabilize. Therefore, agents may wish to continuously update their beliefs about variance to reflect the decreasing variation in the observed market prices (or, equivalently, the decreasing variation in rival firm supplies). This adjustment process can be naturally incorporated into our Bayesian framework, allowing agents to make inferences about *both the mean and the variance* of their rivals' stationary mixed strategies.

#### 4. Bayesian Fictitious Play with Inverse Gamma Prior

In this section, we consider agents making Bayesian inferences about both the mean and the variance of their rivals' mixed strategies using appropriate conjugate priors for these two variables. As before, we examine the oligopoly model with  $M$  firms. Consistent with the fictitious play setup, we assume that each firm  $i$  believes that rival firms are using a stationary random strategy where their aggregate output level  $Q_{-i}$  is sampled from a normal distribution  $Q_{-i,t} \sim \mathcal{N}(\mu_j, \lambda_j)$  with unknown mean  $\mu_j$  and unknown precision (inverse variance)  $\lambda_j$ . Firm  $i$  assumes that the distribution of the hypothesized aggregate output level  $\hat{Q}_{-i}$  of the remaining firms is given by

$$P(Q_{-i}|\mu_i, \lambda_i) = \sqrt{\frac{\lambda_i}{2\pi}} \exp\left(-\frac{\lambda_i}{2}(Q_{-i} - \mu_i)^2\right) \tag{38}$$

The prior beliefs of firm  $i$  about the remaining firms' output choices are given by the conjugate normal-inverse gamma [29] prior on the variables  $\mu_i$  and  $\lambda_i$ :

$$P(\mu_i, \lambda_i | v_{i,0}, \kappa_{i,0}, \alpha_{i,0}, \beta_{i,0}) = \frac{\Gamma(\alpha_{i,0})}{\beta_{i,0}^{\alpha_{i,0}}} \sqrt{\frac{2\pi}{\kappa_{i,0}}} \lambda_i^{\alpha_{i,0}-0.5} \tag{39}$$

$$\exp\left(-\frac{\lambda}{2}[\kappa_{i,0}(\mu_{i,0} - v_{i,0})^2 + 2\beta_{i,0}]\right). \tag{40}$$

The hyperparameters  $v_{i,0}, \kappa_{i,0}, \alpha_{i,0}, \beta_{i,0}$  are known and represent the prior beliefs of firm  $i$  about its rivals' behavior at the start of the Cournot game. As the output choices are revealed in the game, Bayesian posterior estimation is used to update the hyperparameters, using

$$v_{i,t} = \frac{\kappa_{i,t-1}v_{i,t-1} + Q_{-i,t-1}}{\kappa_{i,t} + 1} \tag{41}$$

$$\kappa_{i,t} = \kappa_{i,t-1} + 1 \tag{42}$$

$$\alpha_{i,t} = \alpha_{i,t-1} + 1/2 \tag{43}$$

$$\beta_{i,t} = \beta_{i,t-1} + \frac{\kappa_{i,t-1}(Q_{-i,t-1} - \mu_{i,t-1})^2}{2(\kappa_{i,t-1} + 1)}. \tag{44}$$

The posterior predictive distribution for  $\hat{Q}_{-i}^t$  is given as a Student t-distribution with center  $v_{i,t}$ , precision  $\frac{\alpha_{i,t}\kappa_{i,t}}{\beta_{i,t}(\kappa_{i,t}-1)}$  and  $2\alpha_{i,t}$  degrees of freedom.

$$(\hat{Q}_{-i,t} | \{Q_{-i,s}\}_{s=1:t-1}) = t_{2\alpha_{i,t}}\left(v_{i,t}, \frac{\beta_{i,t}(\kappa_{i,t} - 1)}{\alpha_{i,t}\kappa_{i,t}}\right) \tag{45}$$

As a result, the expected value and variance of  $\hat{Q}_{-i}$  is given by

$$E[Q_{-i,t}] = v_{i,t} \quad V[Q_{-i,t}] = \frac{\beta_{i,t}(\kappa_{i,t} - 1)}{(\alpha_{i,t} - 1)\kappa_{i,t}} \tag{46}$$

The agents choose their output decisions to maximize a linear combination of the expected profit and a term that penalizes variance in the returns.

$$U(q_{i,t}) = E[\pi_i(q_{i,t}, q_{-i,t})] - \gamma V[\pi_i(q_{i,t}, q_{-i,t})] \tag{47}$$

$$= (a - b(q_{i,t} + E[Q_{-i,t}]))q_{i,t} - C_i(q_{i,t}) - \gamma b^2 q_{i,t}^2 V[Q_{-i,t}] \tag{48}$$

$$\tag{49}$$

Therefore, the optimal output choice for the risk-averse firm  $i$  is given by

$$q_{i,t} = \frac{a - d_i - bE[Q_{-i,t}]}{2(b + c_i + \gamma b^2 V[Q_{-i,t}])} \tag{50}$$

Inserting the Bayesian update equations for the expected value and variance of the profit from Equation (46), we obtain

$$q_{i,t} = \frac{a - d_i - bv_{i,t}}{2(b + c_i + \Gamma_{i,t})} \quad \text{with } \Gamma_{i,t} = \gamma b^2 \frac{\beta_{i,t}(\kappa_{i,t} - 1)}{(\alpha_{t,i} - 1)\kappa_{i,t}} \tag{51}$$

This is a nonlinear time-varying difference equation as the terms in  $\Gamma_{i,t}$ ,  $(\beta_{i,t}, \kappa_{i,t}, \alpha_{i,t})$  are *path-dependent*, governed by the realizations of  $q_{j,t}$  for all  $j$ . The results of Theorem 1 extend to this case with asymptotically vanishing variance. Therefore, the firms converge to the Cournot–Nash equilibrium with no offset present in the previous section.

**Theorem 2.** Consider  $M$  firms with identical cost structures  $c_i = c$  and  $d_i = d$  for every firm  $i$ . Then, the dynamical system governed by the state update equations given in (41) and (51) has a fixed point given by

$$q_i^* = \frac{a - d}{2b + (M - 1)b + 2c} \quad v_i^* = \frac{(M - 1)(a - d)}{2b + (M - 1)b + 2c} \tag{52}$$

This fixed point is locally asymptotically stable if  $b/2(b + c) < 1$ . That is, if the initial beliefs are initialized in a small neighborhood of the fixed point, the fictitious play converges to the equilibrium given by the fixed point in the limit as  $t$  approaches infinity.

The proof of the theorem relies on linearizing the nonlinear state evolution equations around the equilibrium point. Let us define the  $2M \times 1$  state vector

$$x_t = \{v_{i,t}, \beta_{i,t}\}_{i=1}^M,$$

that captures all the current mean and variance estimates of each firm. Note that all supply decisions  $q_{i,t}$  are a simple function of the state vector as given in (51) and therefore the convergence of the mean and variance estimates implies the convergence of the output decisions of the Cournot oligopoly. The state evolutions given in (41) and (51) provide a nonlinear update function for the state in the form of

$$x_{t+1} = f(x_t, t)$$

The nonlinear dynamic system is non-autonomous or time-varying due to the simple time dependence of  $\kappa_{i,t} \sim t$  and  $\alpha_{i,t} \sim t/2$ . We observe that it has a fixed point  $x^*$  given by

$$x^* = \left\{ \sum_{j \neq i} q_j^*, 0 \right\}_{i=1}^M,$$

with  $q_i^*$  as defined in (52). In other words, the Bayesian update equations have a fixed point at the Cournot–Nash equilibrium output levels when the variance estimates are equal to zero, an intuitive result. Now, to establish the local asymptotic stability in the neighborhood of the fixed point (for small but non-zero variance estimates), we compute the Jacobian  $J_t$  of the function  $f(x, t)$  around the equilibrium point  $x^*$ . The Jacobian acts as the state transition matrix of the linearized system. By applying the eigenvalue test to the Jacobian  $J_t$ , as shown in the proof of Theorem 1, we obtain the desired result. We note that, in equilibrium, as  $t$  approaches infinity, the term  $\Gamma_t$  asymptotically approaches zero. Consequently, the system converges to the Cournot–Nash equilibrium, with the uncertainty and the associated risk penalty vanishing asymptotically. With risk-averse agents, the stability is improved during the initial stages of learning because non-zero variance estimates and the risk penalty result

in dampening the best response towards zero. The simulation study presented in the next section illustrates this insight.

### 5. Simulation Studies

To study the convergence of the proposed Bayesian fictitious play dynamics numerically, we consider a duopoly model with linear demand and quadratic cost functions. We let the reservation price be  $a = 450$  and the slope of the linear inverse demand function be  $b = 30$ . For this example, we consider identical firms with cost parameters  $c_1 = c_2 = -17$  and  $d_1 = d_2 = 275$ . As a result, the production costs are monotonically increasing, with diminishing marginal costs. We consider output levels in the region such that  $q_{1,t} + q_{2,t} < 15$  to maintain positive prices in the market. We initialize the priors of the agents with the parameters  $\beta_1 = \beta_2 = 1$ ,  $\alpha_1 = \alpha_2 = 5$  and  $\lambda_1 = \lambda_2 = 10$ , with risk aversion factor  $\gamma = 0.01$ . The Cournot game has a unique Nash equilibrium given by  $(2.57; 2.57)$ .

We consider two experiments. In the first experiment, we initialize the beliefs of the agents to  $v_2 = \hat{q}_1 = 4.5$  and  $v_1 = \hat{q}_2 = 0.75$  and the first period output levels as  $q_1 = 4$  and  $q_2 = 1$ , roughly consistent with the prior beliefs. Figure 1a shows the time evolution of the output levels  $q_t$ , as well as the agents' beliefs about one another's output levels  $\bar{q}_t$ . We observe that the output levels and the beliefs converge smoothly to the Nash equilibrium output levels, with  $\Gamma_{i,t}$  converging to zero, capturing a diminishing perceived risk. We also show the output levels corresponding to the traditional best response dynamics with sharper cycles, leading to the same equilibrium level.

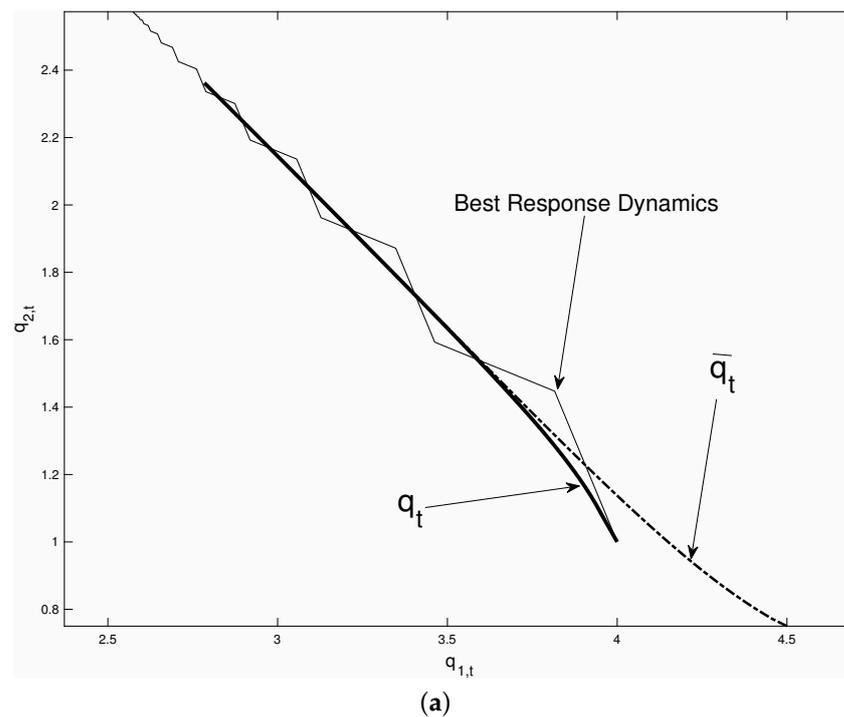
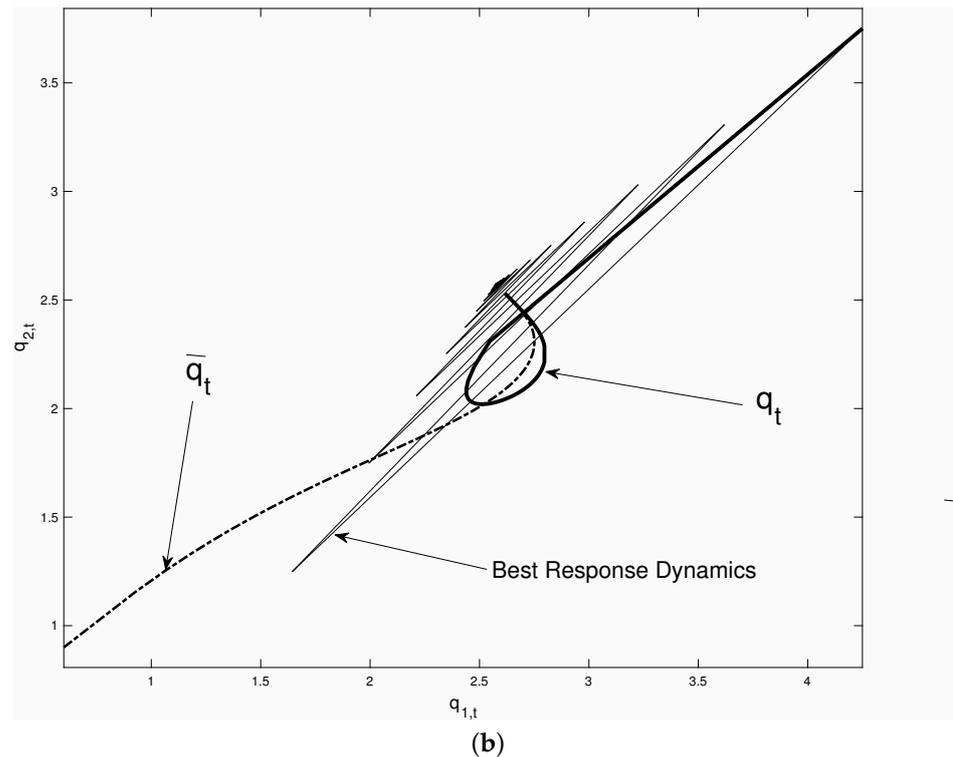


Figure 1. Cont.



**Figure 1.** For the two experiments, the output decisions  $q_t$  of the two firms, as well as their beliefs (mean of the posterior distribution)  $\bar{q}_t$ , converge smoothly to the Nash equilibrium output levels, under the proposed Bayesian fictitious play dynamics. In contrast, the best response dynamics exhibit sharp cycles before convergence.

In the second experiment, we initialize the beliefs of the agents to  $\nu_2 = \hat{q}_1 = 0.6$  and  $\nu_1 = \hat{q}_2 = 0.9$ , inconsistent with the first period's output levels of  $q_1 = 4.25$  and  $q_2 = 3.75$ , modeling a surprise in the observations. Again, Figure 1b shows the time evolution of the output levels  $q_t$ , as well as the agents' beliefs about one another's output levels  $\bar{q}_t$ . We observe that the output levels and the beliefs again converge smoothly to the Nash equilibrium output levels, while the best response dynamics exhibit sharp cycles before eventual convergence.

We note that these simulation results are representative of only oligopoly models with linear demand functions and quadratic cost functions that satisfy the stability criteria given in Theorem 2. Note that when the cost functions are linear (i.e.,  $c = 0$ ), the stability criterion is automatically satisfied and the convergence is guaranteed. We note that the numerical values that we have chosen are consistent with the model used by Cox and Walker [32] in their experimental studies and the model used by Bischi et al. [8] in a study of learning dynamics with risk-neutral agents. This allows a direct comparison to previous experimental and simulation results. Extending this approach to nonlinear demand functions and alternative utility functions that capture varying degrees and forms of risk aversion remains a promising direction for future research.

## 6. Conclusions

In this paper, we present a novel memory-efficient learning model for risk-averse agents in oligopolistic markets. This model extends traditional fictitious play by incorporating Bayesian updates with a normal-inverse gamma prior, enabling firms to learn and adjust both the mean and variance of their competitors' production levels based on the observed market prices. Unlike conventional models, where risk-neutral firms maximize only their expected profits, our approach considers a balance between the expected returns and risk, reflecting real-world complexities where firms face demand and cost uncertainties. We prove that these Bayesian dynamics converge to the Cournot–Nash equilibrium, even

in cases where the best response dynamics would typically diverge, highlighting that risk aversion can stabilize competition in oligopolistic settings.

Through simulation studies, we demonstrate that our Bayesian model achieves smooth convergence to equilibrium, contrasting with the cyclical convergence observed in the best response dynamics. The results emphasize that the initial beliefs impact the convergence behavior but ultimately do not alter the final equilibrium, as Bayesian inference allows firms to adapt based on market data. Our findings offer a robust framework for an understanding of strategic decision-making in oligopolies with risk-averse participants, providing insights that could inform applications in industries where firms face significant uncertainties.

Future research in the area could deepen the understanding of how risk aversion and adaptive learning interact in oligopolistic environments, with potential applications to various real-world industries and market structures. An important extension of our model is to consider inherent uncertainties in the market demand and the ability of the agents to differentiate between persistent demand shocks from the seemingly random behavior of competing agents converging over time to an equilibrium behavior. Alternatively, future work could explore different forms of risk aversion. The current model assumes a linear combination of expected profit and variance penalty; it could be interesting to study alternative forms of risk preference, such as utility-based frameworks or prospect theory, to better reflect the diverse risk attitudes observed in practice. Finally, future work could compare Bayesian fictitious play to other adaptive learning frameworks, such as reinforcement learning or Q-learning, to determine under which conditions each model best approximates the firm behavior and market outcomes in oligopolistic competition.

**Funding:** The research received no external funding.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Conflicts of Interest:** The author declares no conflicts of interest.

## References

1. Fudenberg, D.; Levine, D.K. *The Theory of Learning in Games*; MIT Press: Cambridge, MA, USA, 1998; Volume 2.
2. Theocharis, R.D. On the stability of the Cournot solution on the oligopoly problem. *Rev. Econ. Stud.* **1960**, *27*, 133–134. [[CrossRef](#)]
3. Huck, S.; Normann, H.T.; Oechssler, J. Learning in Cournot oligopoly—An experiment. *Econ. J.* **1999**, *109*, 80–95. [[CrossRef](#)]
4. Huck, S.; Normann, H.T.; Oechssler, J. Stability of the Cournot process—experimental evidence. *Int. J. Game Theory* **2002**, *31*, 123–136. [[CrossRef](#)]
5. Potters, J.; Suetens, S. Oligopoly experiments in the current millennium. In *A Collection of Surveys on Market Experiments*; Wiley Online Library: Hoboken, NJ, USA, 2013; pp. 51–74.
6. Riechmann, T. Cournot or Walras? Long-run results in oligopoly games. *J. Institutional Theor. Econ. (Jite)/Z. Gesamte Staatswiss.* **2006**, *62*, 702–720.
7. Dawid, H. *Adaptive Learning by Genetic Algorithms: Analytical Results and Applications to Economic Models*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2011.
8. Bischi, G.I.; Chiarella, C.; Kopel, M.; Szidarovszky, F. *Nonlinear Oligopolies*; Springer: Berlin/Heidelberg, Germany, 2010.
9. Femminis, G. Risk-aversion heterogeneity and the investment-uncertainty relationship: A closed-form formulation. *Riv. Internazionale Di Sci. Soc.* **2014**, *126*, 275–300.
10. Zhang, C.; Luo, L. Board diversity and risk-taking of family firms: Evidence from China. *Int. Entrep. Manag. J.* **2021**, *17*, 1569–1590. [[CrossRef](#)]
11. Milidonis, A.; Stathopoulos, K. Managerial incentives, risk aversion, and debt. *J. Financ. Quant. Anal.* **2014**, *49*, 453–481. [[CrossRef](#)]
12. Gupta, P. A review of corporate hedging models and their relevance in corporate finance. *Theor. Econ. Lett.* **2017**, *7*, 102–115. [[CrossRef](#)]
13. Asplund, M. Risk-averse firms in oligopoly. *Int. J. Ind. Organ.* **2002**, *20*, 995–1012. [[CrossRef](#)]
14. Kirman, A. On mistaken beliefs and resultant equilibria. In *Individual Forecasting and Aggregate Outcomes*; Cambridge University Press: Cambridge, UK, 1983; pp. 147–166.
15. Kirman, A.P. Learning in oligopoly: Theory, simulation, and experimental evidence. In *Learning and Rationality in Economics*; INSEE: Montrouge, France, 1995; pp. 127–178. Available online: [https://bibliothèque.insee.net/Default/doc/SYRACUSE/538726/learning-in-oligopoly-theory-simulation-and-experimental-evidence?\\_lg=fr-FR](https://bibliothèque.insee.net/Default/doc/SYRACUSE/538726/learning-in-oligopoly-theory-simulation-and-experimental-evidence?_lg=fr-FR) (accessed on 23 November 2024).

16. Brousseau, V.; Kirman, A. The dynamics of learning in N-person games with the wrong N. In *Frontiers of Game Theory*; MIT Press: Cambridge, MA, USA, 1993; pp. 71–93.
17. Bischi, G.I.; Naimzada, A.K.; Sbragia, L. Oligopoly games with local monopolistic approximation. *J. Econ. Behav. Organ.* **2007**, *62*, 371–388. [[CrossRef](#)]
18. Bischi, G.I.; Sbragia, L.; Szidarovszky, F. Learning the demand function in a repeated Cournot oligopoly game. *Int. J. Syst. Sci.* **2008**, *39*, 403–419. [[CrossRef](#)]
19. Brown, G.W. Iterative solution of games by fictitious play. *Act. Anal. Prod. Alloc.* **1951**, *13*, 374–376.
20. Monderer, D.; Shapley, L.S. Fictitious play property for games with identical interests. *J. Econ. Theory* **1996**, *68*, 258–265. [[CrossRef](#)]
21. Alós-Ferrer, C. Cournot versus Walras in dynamic oligopolies with memory. *Int. J. Ind. Organ.* **2004**, *22*, 193–217. [[CrossRef](#)]
22. Bergin, J.; Bernhardt, D. Comparative learning dynamics. *Int. Econ. Rev.* **2004**, *45*, 431–465. [[CrossRef](#)]
23. Vallée, T.; Yıldızoğlu, M. Convergence in the finite Cournot oligopoly with social and individual learning. *J. Econ. Behav. Organ.* **2009**, *72*, 670–690. [[CrossRef](#)]
24. Deschamps, R. An algorithm of game theory applied to the duopoly problem. *Eur. Econ. Rev.* **1975**, *6*, 187–194. [[CrossRef](#)]
25. Dearden, R.; Friedman, N.; Russell, S. Bayesian Q-learning. *Aaai/iaai* **1998**, *1998*, 761–768.
26. Kruse, J.B.; Rassenti, S.; Reynolds, S.S.; Smith, V.L. Bertrand-Edgeworth competition in experimental markets. *Econom. J. Econom. Soc.* **1994**, *62*, 343–371. [[CrossRef](#)]
27. Rassenti, S.; Reynolds, S.S.; Smith, V.L.; Szidarovszky, F. Adaptation and convergence of behavior in repeated experimental Cournot games. *J. Econ. Behav. Organ.* **2000**, *41*, 117–146. [[CrossRef](#)]
28. Thorlund-Petersen, L. Iterative computation of Cournot equilibrium. *Games Econ. Behav.* **1990**, *2*, 61–75. [[CrossRef](#)]
29. DeGroot, M.H. *Optimal Statistical Decisions*; John Wiley & Sons: Hoboken, NJ, USA, 1970; Volume 82.
30. Kirman, A.P. Learning by firms about demand conditions. In *Adaptive Economic Models*; Elsevier: Amsterdam, The Netherlands, 1975; pp. 137–156.
31. Ludyk, G. *Stability of Time-Variant Discrete-Time Systems*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 5.
32. Cox, J.C.; Walker, M. Learning to play Cournot duopoly strategies. *J. Econ. Behav. Organ.* **1998**, *36*, 141–161. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.