Essays on Pricing and Learning in Bertrand Markets

By

Hang Wu

THESIS

Submitted to the University of Adelaide
in partial fulfillment of the
requirement for the degree of

Doctor of Philosophy
in
Economics

October 2013

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution in my name and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works. I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

_____

Signature of Author

Abstract

The thesis studies sellers' pricing and learning behaviour in Bertrand oligopoly markets using a bounded rational approach. It consists of four chapters.

Chapter 1 develops a quantal response adaptive learning model which combines sellers' bounded rationality with adaptive belief learning in order to explain price dispersion and dynamics in laboratory Bertrand markets with perfect information. In the model, sellers hold beliefs about their opponents' strategies and play quantal best responses to these beliefs. After each round, sellers update their beliefs based on the information learned from previous play. Maximum likelihood estimation suggests that when sellers have full past price information, the learning model explains price dispersion within periods and the dynamics across periods. The fit is particularly good if one allows for sellers being risk averse. In contrast, Quantal Response Equilibrium does not organize the data well.

Chapter 2 proposes a generalized payoff assessment learning model of Sarin & Vahid (1999) for the perfect information Bertrand experiments we studied in Chapter 1. The model contains the quantal-response adaptive learning model of Chapter 1 and the original payoff assessment learning model as special cases. A main feature of the model is that it stresses the importance of forgone payoffs for unselected prices in driving the price adjustments. Maximum likelihood estimation shows that the model substantially outperforms the quantal-response adaptive learning model with respect to fitting the data.

Chapter 3 studies the effects of increasing number of sellers on Quantal Response Equilibrium (QRE) prices in homogeneous product Bertrand oligopoly markets. We show that the comparative statics properties of QRE can be very sensitive to the specification of the quantal response function. With the power-function specification, an increase in the number of competing sellers leads to a decrease in the average QRE market price. In stark contrast, with logistic specification, having more sellers may

increase the equilibrium market price, which is at odds with the general intuition that competition should lead to lower prices.

Chapter 4 proposes an extended payoff-assessment learning model to explain the pricing and learning behaviour observed in a repeated Bertrand market experiment with limited feedback. In the experiments, sellers' only feedback after a period was their own payoff. Sellers were not able to observe the prices set by their competitors. The data show that pricing behaviour is strongly influenced by past sales. Sellers on average increase prices after being successful at selling, while they reduce prices after failing to sell. We show that by explicitly incorporating the sellers inferences from the sale history, our learning model manages to explain the data on both the aggregate and individual level.

Acknowledgements

To my mother.

Contents

List of Tables

List of Figures

Chapter 1

Explaining price dispersion and dynamics in laboratory Bertrand markets

1.1   Introduction

For non-economists it is counter-intuitive that in homogeneous product Bertrand markets, in Bertrand-Nash equilibrium, all firms always should set prices at the marginal cost level and earn zero profits. The Bertrand-Nash equilibrium prediction is counter-intuitive to non-game theorists in the sense that no firm has strong incentives to stick to the equilibrium because there is no cost of unilaterally deviating from it. Further to that, sellers are able to achieve higher expected profits by deviating together from the Nash equilibrium and coordinating on higher prices. Observations from real world and laboratory markets cast further doubts on the appropriateness of the Bertrand-Nash prediction. In both real and experimental markets, prices are found to be dispersed above the marginal cost. Moreover, rather than staying constant over time, empirical price distributions show significant inter-temporal variation.

Since Stigler (1961), numerous search-theoretical models have been developed to resolve the puzzle: Salop & Stiglitz (1977), Reinganum (1979), Varian (1980), and Baye & Morgan (2001), to name only a few. By introducing heterogeneity among consumers or sellers in factors such as search costs, production costs or informational frictions, search-

theoretic models provide excellent rationales for price dispersion. Repeated-game effects are also frequently used to explain why prices commonly stay above marginal costs. When sellers engage in repeated interaction, they may have incentives to keep prices at a high level to prevent pricing wars, which might be triggered by a lower price if the competitors adopt a trigger strategy. Alternatively, setting high prices can serve as a signal of friendliness in expectation of reciprocal cooperative behaviour by their opponents in later interactions.

Persistent price dispersion, however, is still commonly recorded in markets where all the aforementioned factors play trivial roles. Evidence can be easily found for online-shopping markets, where search and information costs are negligible (Baye & Morgan 2004) . The same is true for laboratory Bertrand markets where all of those factors can be appropriately controlled for (Bayer & Ke 2011; Dufwenberg & Gneezy 2000). To bridge this gap between theory and empirical reality, some behavioural models have been developed. Rauh (2001) shows that price dispersion can arise when sellers make small but heterogeneous mistakes in beliefs about the market price distribution. Baye & Morgan (2004) show that bounded rational choice models, namely, Quantal Response Equilibrium(QRE, McKelvey & Palfrey 1995) and ε-equilibrium(Radner 1980) can explain price dispersion in homogeneous-good pricing games. Under the QRE model, firms play quantal best responses in a manner that choices with higher expected profits are played with higher probabilities. In contrast, in an ε-equilibrium sellers are equally likely to choose any price that yields an expected profit within ε of the profits for the optimal prices.

Like most of the search-theoretical models of dispersed prices, the bounded rational equilibrium approaches used by Baye & Morgan (2004) are static and hence fail to capture important dynamic features of market prices. In both laboratory and real world price competition markets, market prices typically exhibit significant intertemporal variations. For this reason, we propose a learning model that combines sellers' bounded

2

rational pricing behaviour with learning, which allows for meaningful dynamics.

Our model is based on laboratory observations from two repeated homogeneous product Bertrand experiments. The only difference between the two experiments is the amount of information revealed to the players after each period. More specifically, in one experiment which we call the *high information* treatment, after each period, firms are shown their private profits and all sellers' prices posted in that period. In contrast, in the *low information* treatment, sellers were shown only their own profits. In both treatments, prices are persistently dispersed over the price set. A comparison of the two treatments shows that the information structure influenced sellers' choices substantially. Prices move downward much faster in the high information treatment than in the low information treatment.

We combine sellers' bounded rationality and learning in an attempt to explain price dispersion and dynamics observed in the high information treatment. In our model, the information about past market prices acts as the main factor of driving the price adjustments. Following QRE, we assume that the sellers play quantal best responses to their beliefs about the strategies of their opponents. In an extended model, we increase the flexibility of the model by allowing for different risk preferences. We model the dynamics of the game by a belief learning rule. After each period, based on the previous play, sellers update their beliefs of other sellers' strategies and play quantal responses to the new beliefs. Our model maintains the assumption in Baye & Morgan (2004) that sellers' beliefs take into account the other players' noisy behaviour, which is captured by completely mixed strategies. In QRE, all players' beliefs are consistent with the quantal response choices of their opponents. Rather than considering QRE as an instantaneous result of the game, we conjecture that an equilibrium is a steady state of long-run evolution. We assume that the beliefs of a seller (i.e., a probability distribution over the action space of the opponent) is not necessarily correct and changes according to the learning rule. Following Cheung & Friedman (1997), we use an adaptive learning

3

rule with parameter α measuring how past information is discounted. The learning rule includes Cournot learning ($\alpha = 0$) and fictitious learning ($\alpha = 1$) as extreme cases. When $\alpha \in (0, 1)$, all past interactions affect the beliefs; but the more recent periods receive greater weight.

Note that this model lends itself naturally to one of the treatments (i.e. the full information treatment), while it seems highly inappropriate for the other. Clearly, a model built on sellers learning from past prices in the market place only makes sense if the sellers can observe prices. Consequently, our model provides an appropriate explanation for the price dynamics in Bertrand markets with full information only if it fits well in the high information treatment but at the same time not in the low information treatment. Putting the model to the data of the low information treatment serves as a robustness test. If it were to fit well there, then it could not be ruled out that the potentially good fit in the high information treatment is purely mechanical and results from the number of degrees of freedom in the model.

Maximum likelihood estimates show that the quantal-response adaptive learning model nicely captures the price dispersion and dynamic adjustments observed in the high-information treatment, while the Quantal Response Equilibrium approach of Baye & Morgan (2004) does not. We also find that sellers conditionally on our model being correct exhibit a reasonable degree of risk-aversion. In contrast, our model does not perform well in the the low-information treatment. We conclude that the quantal-response adaptive learning model, where sellers noisily best-respond to their beliefs, is a good explanation for the price dispersion and dynamics in full-information Bertrand markets.

The Chapter is organized as follows. Section 1.2 introduces two Bertrand price competition experiments and its data which we will use as guidance of our modeling. Section 1.3 lays out the QRE model of Baye & Morgan (2004) and our quantal-response adaptive learning model for the finitely repeated Bertrand market game. Section 1.4

uses the experimental data to structurally estimate the parameters of the models, and discusses the results. This Section also conveys a comparison of the goodness of fit for the QRE approach and our learning approach. Section 1.5 concludes with a discussion of the evolutionary properties of the learning model.

1.2   The experiments

In this section, we present two samples of experimental data which will be taken as guidance for our learning model. In section 1.4 we shall also use these data to evaluate the appropriateness of the model. We use truncated data from two 30-period Bertrand price-competition experiments, one from Bayer & Ke (2011) and the other a subsequent experiment, both conducted at Adelaide Laboratory for Experimental Economics (Adlab) at the University of Adelaide. In total, 305 participants participated. The participants were mainly students from the University of Adelaide. They studied for a variety of under and postgraduate degrees. The experiments were designed to investigate the responses of sellers' pricing behaviour to various exogenous cost shocks where the cost shocks were imposed at the beginning of the 16th period. Since the effects of exogenous cost shocks are not of interests for this study, we shall focus on the first 15 periods of play from these experiments for which the production cost were kept constant.

At the beginning of the experiments the participants were randomly assigned roles as sellers or buyers at a fixed ratio of two to one. The roles were kept fixed throughout the experiment. At each period, markets were formed using random matching. Each market consisted of two sellers and a buyer and all subjects were assigned to participate in a market. Random re-matching was adopted to minimize repeated game effects. In each market, two sellers simultaneously and independently set integer prices that

5

could range from E30 (marginal cost) to E100 (reservation value for the buyer).[1] Afterwards, the buyer observes both prices costlessly and then chooses either to buy from one of the sellers or to leave without buying.[2] In each stage the payoff for a seller who managed to sell was her price less the cost. An unsuccessful seller earned a profit of zero. The buyers' payoffs were defined as their reservation value minus the price they paid if they bought and zero if they did not buy. The only difference between the two experiments lies in the information that was revealed to the subjects at the end of each period. In the high information treatment (120 participants), all players were shown their profits and the prices set by **all** sellers in the session[3]. In the low information treatment (185 participants), the participants learned only their own payoffs. No price information were given. Before the experiments, participants were provided with written instructions containing the market rules and the payoff functions. At the end of the experiments, the participants were paid according to their aggregate payoffs in the experimental session. On average they earned around 20 Australian Dollars for about one hour of their time.
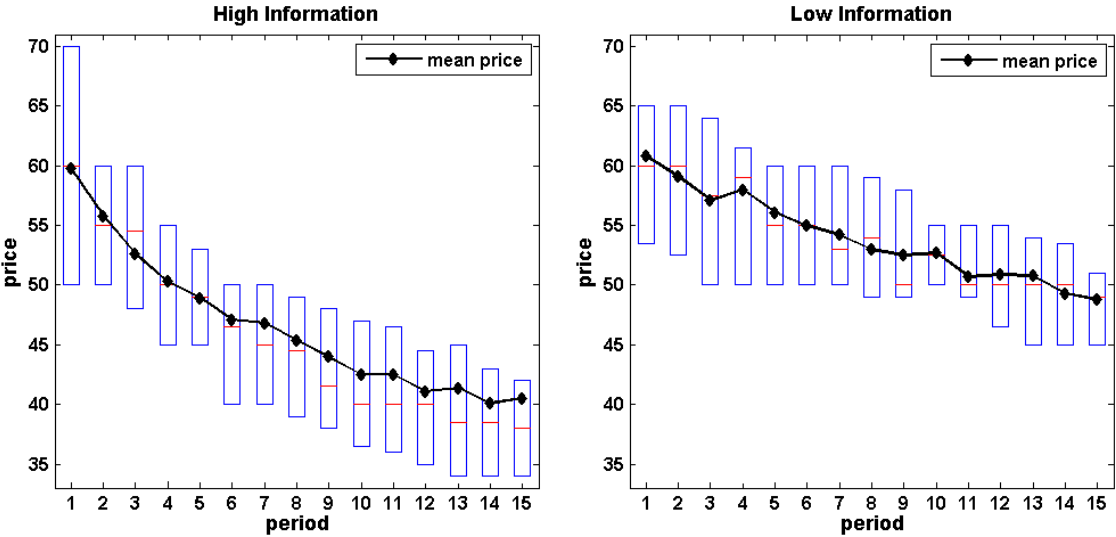
Figure 1.2.1: Time Series for Interquantile Ranges of Prices and Mean Prices

---

[1]The currency was Experimental Dollars. In what follows we drop the currency symbols.

[2]In more than 99% cases the buyers bought from the seller with the lower price.

[3]There were between 12 and 18 sellers in a session.

6

Figure 1.2.1 shows the time series for the interquartile ranges of the prices (boxes) as well as the average prices (black lines). Red bars in the boxes represent the median price levels. For both high information and low information treatments, prices were dispersed persistently over the price set. As can be seen from Figure 1.2.1, in both panels, the central 50 percent of prices exhibit substantial spreads for all periods. In terms of price dynamics, however, starting at virtually identical distributions the prices developed quite differently between the two treatments. For the high information treatment, the average price started off at 59.8, with an interquartile range of 50 to 70 and a median price at 60. Then the prices declined quickly as the experiment proceeded. In period 15, the average price was 40.5, and median price was 38, with an interqartile range of 34 to 42. For the low information treatment, the prices started off at similar levels as in the high information treatment. In period 1, the mean and median prices were 60.5 and 60, respectively, and the corresponding interqartile range was 53.5 to 65. While the prices kept dropping quickly towards Nash equilibrium in the high information treatment, the prices declined at a much slower speed and stabilized in the low information treatment. In period 15, the average price was 48.8, the median price was 49, and the interquartile range 45 to 51. All of these characteristic values are about 10 units above their counterparts in the high information treatment. The fact that the prices evolve significantly differently in the two treatments suggests that feedback on the past strategies plays an important role in price dynamics. In our learning model, the feedback effect will be considered as the main driving force of the price adjustments.

1.3 Theory

1.3.1 Preliminaries

Consider an environment where a set $I = \{1, 2, \cdots, N\}$ of sellers engage repeatedly in the standard Bertrand duopoly game along the time horizon $T \equiv \{1, 2, \cdots, 15\}$. At the beginning of each period, each seller $i \in I$ is randomly matched with a competing seller $j \in I$. Afterwards, seller i and j compete in prices to sell a homogeneous good produced at cost c per unit. The market has unit demand for the good up to a reservation price v. Without loss of generality, we define the price set as $P \equiv [c, v]$. Let (p_i, p_j) be the prices set by the two competing sellers, the payoff to seller i is

$$\pi_i(p_i, p_j) = \begin{cases} p_i - c & \text{if } p_i < p_j \\ \frac{1}{2}(p_i - c) & \text{if } p_i = p_j \\ 0 & \text{otherwise} \end{cases} ; \forall p_i, p_j \in P. \qquad (1.3.1)$$

Let seller i's strategy be a cumulative probability measure over the price set, denoted as $F_i : P \to [0, 1]$. Further, let $B_i(F_j)$ be seller i's belief about her rival j's strategy. Thus, the expected monetary payoff for seller i posting price p, given $B_i(F_j)$, is

$$E\pi_i(p) = (p - c)\left[1 - B_i(F_j(p))\right], \forall i, j \in I, i \neq j. \qquad (1.3.2)$$

We now state, without proof, the well known Bertrand-Nash equilibrium, where all probability mass in a mixed strategy is put on the price that equals marginal cost.

Proposition 1.1. *(Bertrand-Nash equilibrium) For all periods $t \in T$, the following comprises a symmetric Bertrand-Nash equilibrium: For all $i, j \in I$ and for all $p \in P$, $F_i^{NE}(p) = 1$.*

1.3.2 Quantal Response Choices

The first extension to the Nash equilibrium tradition in our model is to introduce bounded rationality to sellers' pricing decisions. It is assumed that sellers are prone to choice errors and post suboptimal prices with positive probabilities. The errors could be caused by inexperience, computational limits, or instantaneous mood shocks (see Chen et al. 1997). Following Baye & Morgan (2004) and López-Acevedo (1997), we incorporate the choice errors using a power-form quantal-response function. Formally, the strategy of seller $i \in I$ in terms of cumulative probability distribution is

$$F_i(p, B_i) = \frac{\int_{q=c}^{p} \left[E\pi_i(q, B_i(F_j))\right]^\lambda dq}{\int_{k=c}^{v} \left[E\pi_i(k, B_i(F_j))\right]^\lambda dk}, \ \forall p \in \mathrm{P} \tag{1.3.3}$$

where

$$E\pi_i(p, B_i(F_j)) = (p - c) \left[1 - B_i\left(F_j(p)\right)\right]. \tag{1.3.4}$$

The probability that seller i set her price at p is positively related to the expected monetary payoff of p. The "error parameter" or "bounded-rationality parameter", $\lambda \in [0, \infty)$, measures the degree of sensitivity of the firms to the expected payoffs. As $\lambda \to \infty$, the firm tends to choose the payoff maximizing price with certainty and becomes fully rational. On the other hand, as $\lambda \to 0$, the firm becomes fully ignorant or confused and randomizes over all prices with equal probabilities. A property of the choice-probability functions that is important for empirical applications and also for economic intuition is that strategies that yield greater expected payoffs are chosen with higher probabilities when $\lambda > 0$. A noteworthy special case is $\lambda = 1$, under which choice probabilities of firm i are proportional to the expected payoffs. This is the classic Luce (1959) probabilistic choice model and was first applied in non-cooperative games by Rosenthal (1989).

1.3.3 Quantal Response Equilibrium (QRE)

In a Quantal Response Equilibrium (QRE), all sellers' strategies are quantal responses to their beliefs about the probability distributions of their opponents' prices. That is, for each $i \in I$, F_i follows Equation 1.3.3. Moreover, the beliefs of all sellers are consistent with the probability distributions of their opponents' prices. We have $B_i(F_j) = F_j, \forall i, j \in I$. Baye & Morgan (2004) obtained a simple, closed-form representation of QRE pricing strategies for the homogeneous product Bertrand duopoly game.

Proposition 1.2. *(Baye & Morgan 2004) For any $\lambda \in [0, 1)$, the following comprises a symmetric QRE:*

$$F_i^Q(p) = 1 - \left[\frac{\pi(v)^{1+\lambda} - \pi(p)^{1+\lambda}}{\pi(v)^{1+\lambda} - \pi(c)^{1+\lambda}} \right]^{\frac{1}{1-\lambda}} \forall p \in [c, v], \ \forall i \in I \qquad (1.3.5)$$

where $\pi(p)$ is the payoff to a monopolist charging price p.

Proof. See Baye & Morgan (2004).

$\square$

Note that for the QRE to exist we must have $\lambda \in [0, 1)$. When λ is greater than one, according to Equation (1.3.5), we have $F_i^Q(p) > 1$ for all prices that are between c and v, which is impossible. With $\lambda = 0$, the sellers behave randomly in choosing prices so that the prices are distributed uniformly over the price set P. At the other extreme, with λ tends to 1, more and more probability mass is allocated to low prices and the QRE converges to the Bertrand-Nash equilibrium. Therefore, in this game, to attain the Bertrand-Nash equilibrium result, perfect rationality ($\lambda \to \infty$) is not required.

An appealing feature of the quantal-response choice rule is its flexibility that allows for incorporating and parametrization of factors that may influence players' behaviour other than bounded rationality. By using the Arrow-Pratt risk measure, we can extend the QRE model to allow for heterogeneous attitudes toward risk or uncertainty in different circumstances of the game.[4] Formally, instead of maximizing the expected monetary payoffs, we assume that the sellers aim to maximize expected utilities which we define as

$$EU_i(p) = \frac{(p-c)^{1-r}}{1-r}[1 - B_i(F_j(p))]; \forall t \in T, \forall i, j \in I. \tag{1.3.6}$$

The parameter r measures a seller's risk attitudes, with $r = 0$ corresponding to risk neutrality, $r > 0$ to risk aversion, and $r < 0$ to risk seeking.[5]

Proposition 1.3. *(QRE with Arrow-Pratt Risk Attitudes) For any $\lambda \in [0, 1)$ and $r < 1$, the following comprises a symmetric QRE with risk attitudes:*

$$F_i^{AP}(p) = 1 - \left[1 - \left(\frac{p-c}{v-c}\right)^{(1-r)(1+\lambda)}\right]^{\frac{1}{1-\lambda}}. \tag{1.3.7}$$

Proof. Setting $\pi(p) \equiv \frac{(p-c)^{1-r}}{1-r}$ in Equation 1.3.5 yields the result.

$\square$

---

[4]See Goeree et al. (2002) for an example that incorporates QRE with risk aversion in explaining overbidding in private value auctions.

[5] The utility function we use exhibits constant relative risk aversion and is used frequently in experimental research (e.g., Holt & Laury 2002). For $r = 1$, where the expected utility function is undefined, we use $\ln(p-c)$ instead of $\frac{(p-c)^{1-r}}{1-r}$. This is because for $r \to 1$ we have $\frac{d\ln(x)}{dx} = \frac{d(\frac{x^{1-r}}{1-r})}{dx}$.

1.3.4 A Quantal Response Adaptive Learning (QRAL) model

In this subsection we propose a simple quantal response learning model to explain sellers' intertemporal price adjustments observed in the high information treatment. It is assumed that firms formulate beliefs of their competitiors' future strategies based on the price information of the past periods and play quantal responses to their beliefs. We use an approach similar to the empirical learning rule of Cheung & Friedman (1997) where a player's belief is the weighted average of the strategies that she encountered in the past periods. Cournot learning and fictitious learning are special cases of the model. While Cheung & Friedman (1997) assume that players' beliefs are formulated using the past strategies of their actual rivals, we assume that the sellers' current beliefs to be the weighted average of all her potential opponents' past strategies. This is a reasonable assumption because in our context the sellers are randomly rematched in each period and are shown the prices of all sellers in the same session. Let $(F_1, \cdots, F_t)$ denote the vector of market price distribution observed from period 1 to period t , the belief firm i holds before period $t + 1$ is:

$$B_{i,t+1}(F_{j,t+1}) = \frac{F_t + \sum_{\tau=1}^{t-1} \alpha^\tau F_{t-\tau}}{1 + \sum_{\tau=1}^{t-1} \alpha^\tau}. \tag{1.3.8}$$

Parameter α captures the feature that different past histories enter with different weights into the beliefs. When $0 < \alpha < 1$ we have the typical case that recent histories carry more weight than older histories. Setting $\alpha = 0$ yields the Cournot adjustment rule, where only the most recent period is relevant for the beliefs. Setting $\alpha = 1$ yields standard fictitious play, where all past experiences are weighed evenly. Consequently, in period $t + 1$ seller i's probability choice function, given her belief $B_{i,t+1}(F_{j,t+1})$, can be written as

$$F_{i,t+1}(p) = \frac{\int_{q=c}^{p} \left[EU_i\left(q, B_{i,t+1}(F_{j,t+1})\right)\right]^\lambda dq}{\int_{k=c}^{v} \left[EU_i\left(k, B_{i,t+1}(F_{j,t+1})\right)\right]^\lambda dk}. \tag{1.3.9}$$

1.4 Estimation

In this section we use our experimental data to estimate the parameters of the QRE and QRAL models and evaluate the relative fit of these models. We use a discretized version of the quantal response function and adopt the following *interiority* condition (cf. Goeree et al. 2005):

$$f_{i,t}(p, B_{i,t}) = \frac{\left[EU_i\left(p, B_{i,t}(F_{j,t})\right)\right]^{\lambda}}{\sum_{k=c}^{v}\left[EU_i\left(k, B_{i,t}(F_{j,t})\right)\right]^{\lambda}} > 0, \forall i, j \in I, \forall t \in T, \forall p \in \mathrm{P}.$$

That is, the mixed strategies defined by the quantal response functions are complete so that all prices in P are played with positive probabilities. An example in which the *interiority* condition is violated is $B_i(F_j(p)) = 1$ for all p in P. In this case, player i believes that c is played with certainty by firm j, so the expected payoff for any price is zero. Hence, both the numerator and denominator of the quantal response function are equal to zero, which causes an indeterminacy problem. To avoid such indeterminacy and to ensure that the *interiority* condition is satisfied, in our estimations, we adjust the expected payoffs by adding a small positive technical parameter ε:

$$EU_{i,t}(p, B_{i,t}) = \varepsilon + \frac{(p-c)^{1-r}}{1-r}\left[1 - B_{i,t}\left(F_{j,t}(p)\right)\right]; \forall t \in T, \ \forall i, j \in I \qquad (1.4.1)$$

One justification for ε is that when people take part in economic activities, they receive some level of satisfaction from participating, which is independent of the monetary outcomes they get from the activities. For example, in most economics experiments, subjects are rewarded with a show-up fee for participation in addition to earnings that are proportional to their performance. The introduction of ε considerably facilitates the empirical application of the model. Also, when ε is sufficiently small, compared to the general expected payoffs, it will not change any of the main implications of the model. After the transformation, when facing $B_{i,t}(F_{j,t}(p)) = 1, \forall p \in \mathrm{P}$, the quantal

response choice function will assign uniform probabilities to all prices in P, which is intuitively and economically more convincing because we naturally expect prices with identical payoffs to carry the same weight in the sellers strategies.[6]

We use maximum likelihood estimation (MLE) to derive the estimates for the parameters of interest. To do this, we search numerically for the parameters that maximize the likelihood of occurrence of the set of prices observed in the experiments. We set $\varepsilon = 10^{-10}$ and take the first period's price distribution as the sellers' initial beliefs. From period 2 to 15, for each treatment we calculate the probabilities associated with the prices chosen by sellers using the quantal response function. The log-likelihood function is

$$\log(L) = \log \left[\prod_{t=2}^{15} \prod_{i=1}^{N} f_{i,t}(p_{i,t}, B_{i,t}) \right] = \sum_{t=2}^{15} \sum_{i=1}^{N} \log \left[f_{i,t}(p_{i,t}, B_{i,t}) \right]; \qquad (1.4.2)$$

where N is the number of sellers participating in a treatment. For the purpose of comparison, we conduct the ML estimations for both the QRE and QRAL models in both high and low information treatments. Recall that the low information treatment is inconsistent with our learning model as the seller does not have the information required. The fit in this situation, where the model is misspecified by design, will be used as a robustness test.

---

[6] In our context, the adoption of the technical parameter ε only serves to prevent the indeterminacy problem that may arise in the power function specification of the quantal response function. We use the power specification to keep our model in line with the Baye & Morgan (2004) approach. Alternatively, we can use the logistic specification, which allows for zero and negative payoffs.

	risk neutral($r = 0$)				risk seeking/averse ($r \neq 0$)				
	λ	α	$\log(L)$	BIC	λ	α	r	$\log(L)$	BIC
QRAL	1.542	0.321	-3883	3890	1.522	0.505	0.365	-3826	3837
($\alpha \in [0, 1]$)	(0.021)	(0.010)			(0.014)	(0.029)	(0.017)		
QRE	0.914	–	-4122	4125	0.842	–	0.255	-4027	4034
	(0.022)				(0.011)		(0.003)		

Table 1.4.1: Maximum Likelihood Estimates for the High Information Treatment.

Table 1.4.1 reports the maximum likelihood estimates for the high information treatment. In parentheses are standard errors obtained using numerical differentiation. The table also includes the log-likelihood value $\log(L)$ and the Bayesian Information Criterion (BIC), which is used to compare the relative goodness of different models.[7] According to BIC, QRAL outperforms QRE in a substantial way. Adding a risk-preference parameter improves the fit even when the additional degree of freedom is taken into account. In the learning model with risk preference parameter, the decay parameter $\hat{\alpha}$ is estimated as 0.505, which is significantly different from both zero and one. This is reasonable because we would naturally expect the influence of past history to decay as the experiment proceeds. With $\hat{\alpha} = 0.5$, the price information from more than four periods ago has lost almost all of its influence on today's beliefs (as $0.5^5 \approx 0.03$). For both QRE and QRAL models, the parameter $\hat{r}$ is positive, which indicates that sellers' choices were guided by risk aversion.[8] For the QRE models, the estimates for the bounded rationality parameter $\hat{\lambda}$ are equal to 0.914 and 0.842, respectively for the models with and without risk parameter. In contrast, if we allow for learning, both $\hat{\lambda}$'s are about 1.5. This suggests that with QRAL the price dynamics may fail to converge

[7] In our analysis, BIC is defined as $BIC = \frac{k}{2}\ln(N*T) - \ln(L)$. Here k is the number of parameters, N is the number of sellers, T is the number of periods considered, and $\ln(L)$ the value of the log-likelihood function. BIC penalizes models with additional parameters. According to this criterion, a model with lower BIC value is preferred.

[8] When $\hat{r} = 0.365$, a subject is willing to pay about 35 dollars to take a gamble that yields zero and 100 with the same probability 0.5.

to a QRE because for QRE to exist, as indicated in subsection (1.3.3), λ needs to be less than one. So far we can conclude that a model that combines a belief learning with quantal best response behaviour dominates the static QRE model typically used to explain pricing behaviour in Bertrand duopolies. Adding a risk parameter further improves the explanatory power of the model.
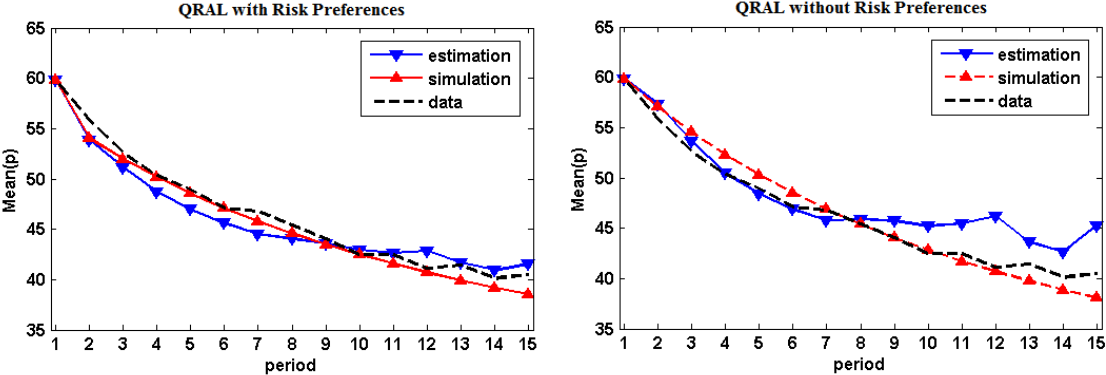
Figure 1.4.1: Time Series of Mean Prices: Data and QRAL Predictions-High Information Treatment

Figure 1.4.1 plots the estimated mean prices for the adaptive learning models, along with the empirical mean prices for the high-information treatment. It also shows the mean prices of simulations using the estimated parameters. In the simulations, we adopt market-price distribution of period 1 as the initial belief, $B_{i,2}(F_{j,2}) = F_1$. Then given the values of $\hat{r}$, $\hat{\alpha}$ and $\hat{\lambda}$, we can obtain the predicted mixed strategies for period 2. Then instead of using the actual observed price distributions, the simulations use the predicted strategies to form the new beliefs and proceed by iterating forward on the system to obtain the simulated strategies for all periods. Therefore, the difference between the simulations and estimations is that in the simulations we use the strategies predicted by Equation 1.3.9 to formulate the beliefs, while in the estimations we use actually documented strategies. As can be seen from Figure 1.4.1, the average prices predicted by the QRAL model are fairly close to the empirically observed dynamics. In particular, in our preferred model (i.e. QRAL with risk preferences) all three time

series – empirical, estimated and simulated prices – are very close together.
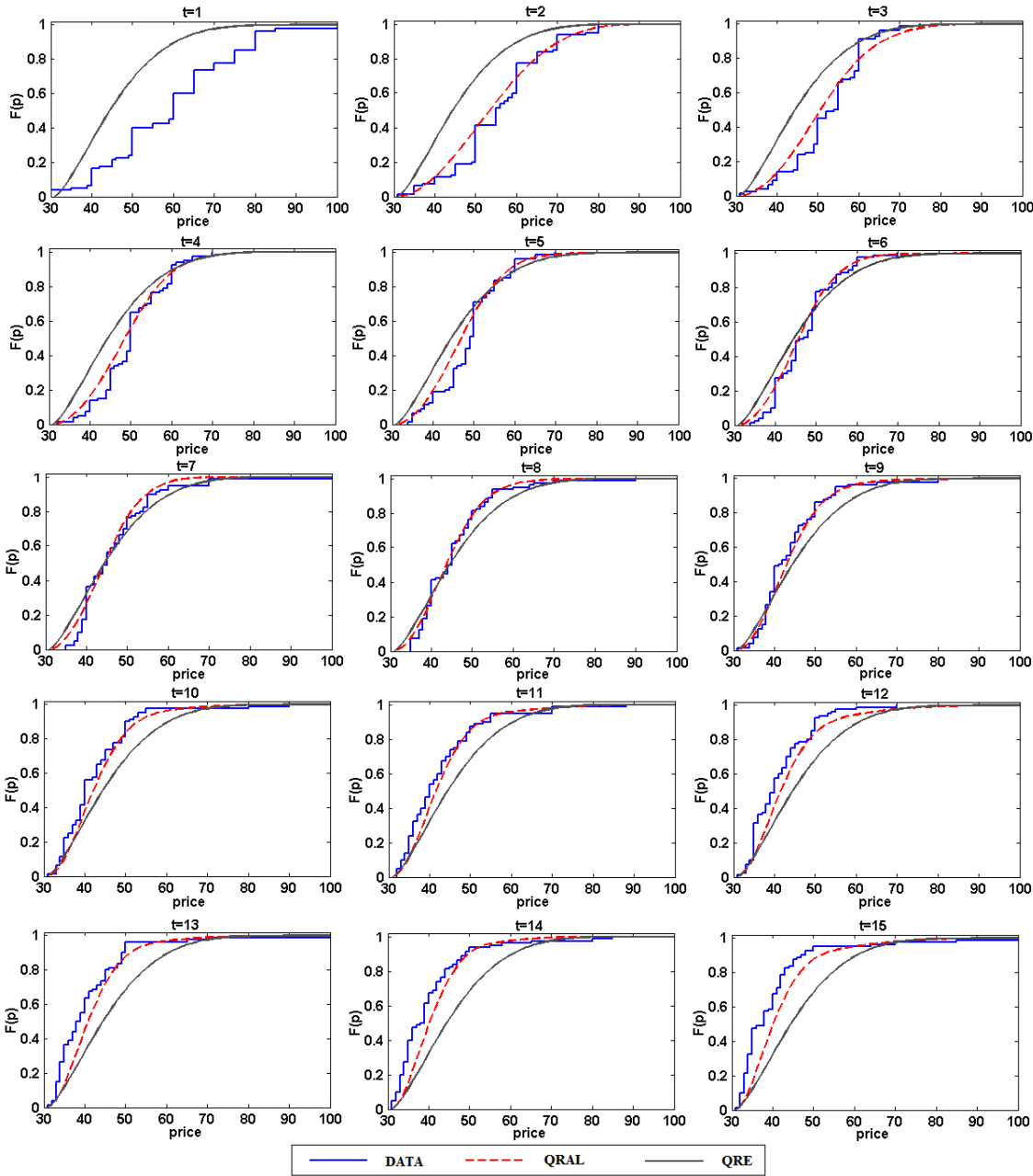
Figure 1.4.2: Comparisons of Cumulative Distributions of Prices: Data, QRAL with Risk Preferences, and QRE with Risk Preferences

Figure 1.4.2 shows the estimated distributions of prices predicted by the QRE and QRAL models separately for the 15 periods, both with risk-preference parameters, along with the empirical price distributions observed in the experiments. The plots indicate

that the QRAL model predicts the price dispersion and and its evolution quite well, while the QRE model, due to its static nature, works only well for some middle periods and fails to capture the price distribution dynamics.

	risk neutral($r = 0$)				risk seeking/averse ($r \neq 0$)				
	λ	α	$\log(L)$	BIC	λ	α	r	$\log(L)$	BIC
QRAL	1.578	0.865	-6176	6183	1.436	0.476	-0.602	-6101	6112
($\alpha \in [0, 1]$)	(0.021)	(0.016)			(0.016)	(0.009)	(0.010)		
QRE	0.836	–	-6411	6415	0.894	–	-0.235	-6347	6354
	(0.011)				(0.021)		(0.011)		

Table 1.4.2: Maximum Likelihood Estimates for the Low Information Treatment.
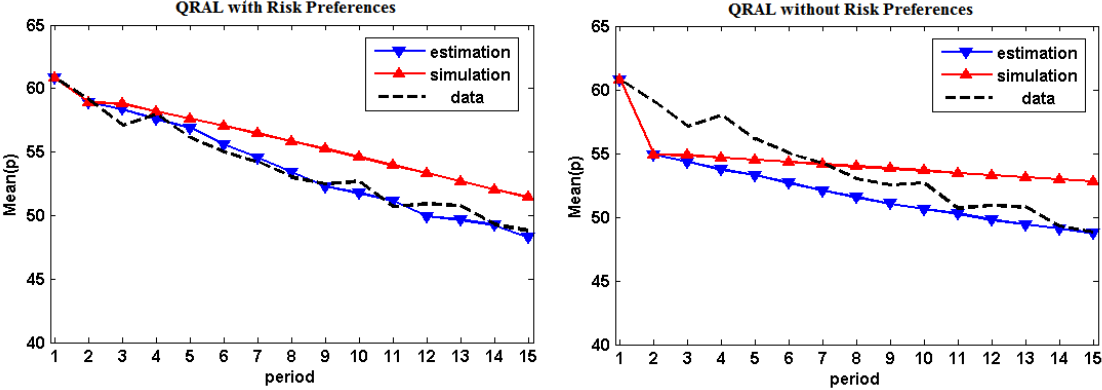
Figure 1.4.3: Time Series of Mean Prices:Data and QRAL Predictions-Low Information Treatment

When we apply the same methodology to the low information treatment data we find that the estimated parameter values are not plausible and the model fit is poor. Table 1.4.2 shows our estimation results for the different models. In order to achieve a reasonable fit we require an unreasonably low risk-preference parameter r of less than -0.6. In the otherwise best-fitting QRAL model, this implies an unreasonably high level of risk-love.[9] Moreover, it becomes clear that, when we plot estimated and simulated price time series against the observed prices (Figure 1.4.3), that the model

---

[9]A person with such risk preferences prefers a gamble of $2 with probability 1/3 and nothing with probability 2/3 to receiving $1 for certain.

does a poor job at explaining the pricing behaviour in the low-information treatment. The poor performance of the model in the low information treatment indicates that the model's good fit in the high information treatment is not merely an artifact of its degrees of freedom. Consequently, we conclude that the estimated QRAL model by mildly risk-averse subjects is a robust explanation for observed behaviour in Bertrand duopolies with perfect past information.

1.5   Concluding remarks

In this Chapter we developed a quantal response adaptive learning model in order to explain price dispersion and dynamic adjustments observed in repeated experimental Bertrand markets, where prices are observable. In our model, rather than being fully rational and choosing only crisp best responses, the sellers are assumed to be boundedly rational in the sense that they play suboptimal strategies with positive probabilities. The probability they play a specific strategy with is a monotonic function of that strategy's expected payoff. We show that price dispersion can be effectively explained by such quantal response choice rules. However, the static equilibrium approach based on the quantal-response choice rule, QRE, fails to explain the evolution of prices over time. We use an adaptive belief-learning rule to model learning and to explain price dynamics. The beliefs of the sellers are assumed to be the weighted average of her rivals' past strategies. We show that for experiments, where the sellers have perfect information on their opponents' past choices, the quantal response adaptive learning model can explain the dynamic evolution of prices remarkably well. In contrast, the model fails to provide reasonable estimates for experiments where no price information of preceding play was revealed. This result has two important implications. Firstly, the good fit of the adaptive learning model is not an artifact just stemming form the model's

degrees of freedom. Secondly, an alternative model with limited past information is necessary for the low-information treatment.

We want to conclude this Chapter with an out-of sample investigation. It is interesting to study how the price dynamics of our best model will evolve if we extend our investigation to a time horizon that is longer than what we had in the experiments. We simulate the QRAL model using the estimated parameters for 200 periods. The result shows that the average price will evolve cyclically without stabilizing at an equilibrium. Figure 1.5.1 shows the evolution of the average price. The intuition behind the cyclicality is the following: whenever the mass of the price distribution gets pushed towards marginal cost, then the profitability of charging a price close to marginal cost is very low. Thus it becomes profitable to charge a higher price and hope for the rare occurrence of a competitor who charges a high price due to bounded rationality. Given the adaptive nature of the learning process, many sellers will follow this strategy at the same time. Thereafter the downwards dynamics sets in again until a jump becomes profitable again.[10] In our experiments, we do not have enough time periods in order to see if the cyclical pattern emerges. Bruttel (2009) found some cyclical movement in her series of experiments, which provides some evidence.

---

[10]Note that the cyclicality requires a sufficient level of rationality (i.e. a λ of greater than one) to occur.
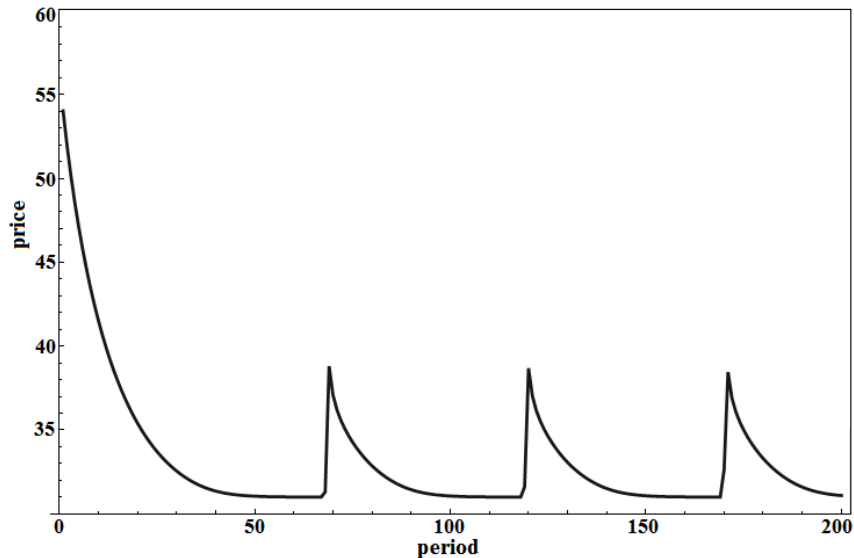
Figure 1.5.1: Simulation of the QRAL model with estimated parameters for 200 periods

Theoretical explanations of price cycles have focused on the Edgeworth cycles. Maskin & Tirole (1988) show that in dynamic Bertrand duopoly games, if sellers follow alternating-move dynamics and adopt Markov perfect equilibrium, then cyclical prices will be a natural result. In their model, sellers engage in price undercutting until they arrive at a bottom price, at which the equilibrium strategy for the firm who gets to move is to raise prices with positive probability. When the firm raises its price, a new price cycle is triggered. Our model provides an alternative explanation for the cyclical price phenomenon in Betrand markets. As opposed to the alternating-move assumption that only one firm gets to move in each period, we allow both firms to adjust their prices in all periods.

Chen et al. (1997) show analytically that for any finite game where the payoffs are positive for all players, if the choices of players are noisy enough, or put equivalently, if the bounded rationality parameter λ is small enough, then fictitious play converges to a unique noisy learning equilibrium. For our adaptive learning model, conditions in which the prices converge and in which the dynamics fail to converge still need to be

investigated formally. Moreover, a more in-depth experimental investigation is required in order to test if price-cycles occurring in laboratory studies are consistent with our theory.

Chapter 2

Price dispersion and dynamics: A payoff assessment learning approach

2.1   Introduction

In Chapter 1 we used a quantal-response adaptive learning model to explain price dispersion and learning behaviour observed in our Bertrand duopoly experiments with perfect information. In this Chapter we propose an alternative model to address the same issue. The model is an extension of the Payoff Assessment Learning (PAL) model of Sarin & Vahid (1999). PAL was designed for limited-information settings where the only feedback a player receives after a period is her own payoff. We generalize PAL and make it suitable for the use in perfect-information settings. Our model contains the quantal-response adaptive learning model (from Chapter 1) and the original PAL model of Sarin & Vahid (1999) as special cases. We then apply the model to the same data set of high-information Bertrand duopoly experiments used in Chapter 1. Maximum likelihood estimation shows that the model substantially outperforms the quantal-response adaptive learning model with respect to fitting the data.

The next Section lays out the model. Section 2.3 presents the estimation and Section 2.4 concludes.

2.2 A generalized payoff assessment learning model (GPAL)

Players repeatedly play the Bertrand duopoly game we introduced in Chapter 1. A set of sellers $I = \{1, 2, \cdots, N\}$ play the stage game for T periods. In each period $t \in \{1, 2, \cdots, T\}$, a seller is competing against a randomly determined competitor. Denote the price set as $P = \{c, c+1, \cdots, v\}$. The payoff for seller i is $p_i - c$ if her price is lower than that of her opponent, is $\frac{p_i - c}{2}$ in the case of a tie, and zero otherwise.

Let $A_t^i(p)$ denote seller i's payoff assessment of choosing price p in period t. After having observed the information, the law of motion for seller i's payoff assessment of price p is governed by:

$$A_{t+1}^i(p) = \alpha A_t^i(p) + \beta U_t^i(p) + (1 - \alpha - \beta)V_t^i(p); \ \forall i \in I, \ \forall p \in P. \tag{2.2.1}$$

Here $A_t^i(p)$ is seller i's payoff assessment for price p in the previous period, which we refer to as the ***inertia factor*** of learning. The second component in Equation 2.2.1 $U_t^i(p)$, called the ***belief-learning factor***, is seller i's expected payoff if she had chosen p in response to the realized market price distribution. Denote the cumulative probability distribution function for price $p \in P$ in period t as $F_t(p)$. The corresponding density function is written as $f_t(p)$. We can calculate the belief-learning factor as $U_t^i(p) \equiv (p-c)[1 - F_t(p) + \frac{1}{2}f_t(p)]$. The remaining component $V_t^i(p)$ is the ***experiential-learning factor***. More specifically, $V_t^i(p)$ is the (forgone) payoff to seller i if she had posted p while facing the same opponent she actually faced in the period before. The parameters $\alpha \in [0, 1]$ and $\beta \in [0, 1]$ measure the weights that the sellers allocate to $A_t^i(p)$ and $U_t^i(p)$, respectively; and $1 - \alpha - \beta$ is the weight being allocated to $V_t^i(p)$. Note that, if $1 - \alpha - \beta = 0$ the model becomes the risk-neutral adaptive belief learning model (see Chapter 1) where sellers play in response to the weighted sum of all past periods' market price distributions. If on the other hand we have $\beta = 0$, the model is

turned into a modified payoff assessment learning model.[1] Also noteworthy here are the two cases corresponding to the two extreme values of α. With $\alpha = 0$ a seller's payoff assessments are depending purely on the information about the most recent play, and with $\alpha = 1$ the market is at a steady state and the sellers always stick to their initial strategies.

Given the payoff assessments $A_t^i(p)$ of each seller $i \in N$ for each price $p \in \mathrm{P}$, we can transform the payoff assessments into the choice probabilities by either the logisitic choice rule as

$$f_{t+1}^i(p) = \frac{e^{\lambda \cdot A_t^i(p)}}{\sum_{k=c}^{v} e^{\lambda \cdot A_t^i(k)}};$$
(2.2.2)

or the power form choice rule as

$$f_{t+1}^i(p) = \frac{\left[A_{t+1}^i(p)\right]^\lambda}{\sum_{k=c}^{v} \left[A_{t+1}^i(k)\right]^\lambda}.$$
(2.2.3)

2.3   Estimation and Results

In this section we apply the model to the high-information Bertrand duopoly experiments where all elements of the information stated in Equation 2.2.1 are available to the sellers.[2] More specifically, after a period a seller was provided with the following pieces of information: her realized payoff, her actual opponent's price, and the prices posted by all sellers in the same session of the experiments. Let p_t^j denote seller i's actual opponent's price in period t. Then the experiential-learning component $V_t^i(p)$

---

[1] The original payoff assessment learning model of Sarin & Vahid (1999) considers only the realized payoffs for the chosen actions. Our model takes into account the forgone payoffs for the unchosen actions. When a player can learn their actual opponents' action, the forgone payoff for an alternative action can be easily calculated. As we shall show in Chapter 4, even if the information is very limited, a player might still be able to infer the forgone payoffs for the unchosen actions.

[2] For the details of the experiments, please refer to Chapter 1.

can be stated as:

$$V_t^i(p) = \begin{cases} p - c & \text{if } p < p_t^j \\ \frac{1}{2}(p - c) & \text{if } p = p_t^j \\ 0 & \text{otherwise} \end{cases} \qquad (2.3.1)$$

Now, we turn to the estimation of the parameters. We take the observation in the first period as a starting point, the initial assessments of prices, $A_1^i(p) \, \forall p \in P, \, \forall i \in I$, is set to the expected payoff of p being played against the market price distribution in period 1. Starting with $A_1^i(p)$, when new information is revealed, then the payoff assessment vector is updated according to Equation 2.2.1, and the choice probabilities can be determined by Equation 2.2.2 or Equation 2.2.3. We then search for the values of α, β and λ that maximize the log-likelihood of observing the prices in the experiments from period two to fifteen. To allow for a direct comparison between GPAL and the model used in Chapter 1, the power-form probabilistic choice rule is adopted.[3] So, the estimation task can be stated formally as

$$\max_{\alpha,\beta,\lambda} \log(L) = \log\left[\prod_{t=2}^{15}\prod_{i=1}^{80} f_t^i(p_t^i)\right] = \sum_{t=2}^{15}\sum_{i=1}^{80} \log\left[\frac{[A_t^i(p)]^\lambda}{\sum_{k=c}^v [A_t^i(k)]^\lambda}\right]. \qquad (2.3.2)$$

α	β	λ	$\log(L)$	BIC
0.518	0.023	1.522	-3457	3467
(0.032)	(0.007)	(0.055)		

Table 2.3.1: Maximum likelihood estimation-Results

Table 2.3.1 presents the results for the estimation. In the parenthesis are the standard errors calculated by numerical differentiation. The best value of $\beta = 0.023$ indicates that the belief-learning component has a very small but significant effect on the

---

[3]We have done the estimation using both choice rules, the power form resulted in a better fit than the logistic form in the estimation.

price adjustments. In contrast, the inertia component and the experiential component of payoff assessment are the main driving force of price adjustment. The estimate of $\lambda = 1.522$ is very close to the value we derived from the quantal response adaptive learning (QRAL) model in Chapter 1. Recall that in the risk-neutral QRAL model we have $\lambda = 1.542$, and in the QRAL risk-nonneutral model we have $\lambda = 1.522$, which is exactly the same as what we derive here. The log-likelihood value -3457 and Bayesian Information Criterion (BIC) value 3467 indicates that GPAL organizes the data much better than QRAL in terms of goodness of fit. [4]

Figure 2.3.1 shows observed data and the model predictions for the average prices. Figure 2.3.2 shows the corresponding variances of prices. In addition, Figure 2.3.3 provides a comparison of the actual and predicted relative frequencies from period two to period fifteen. As can be seen from the figures, GPAL organizes the data rather well. To evaluate the performance of the model in a more convincing way, we conduct Kolmogorov-Smirnov(K-S) tests to test the null hypothesis of equality between the actual price distributions and those predicted by the model. For 10 out of 14 periods (71.4%) the tests fail to reject the null hypothesis, which provides strong evidence for the good performance of GPAL. .

---

[4]According to the Bayesian Information Criterion, a model with lower BIC is preferred. Recall that for the QRAL, the best fitting model has BIC=3837, which is much higher than 3467.
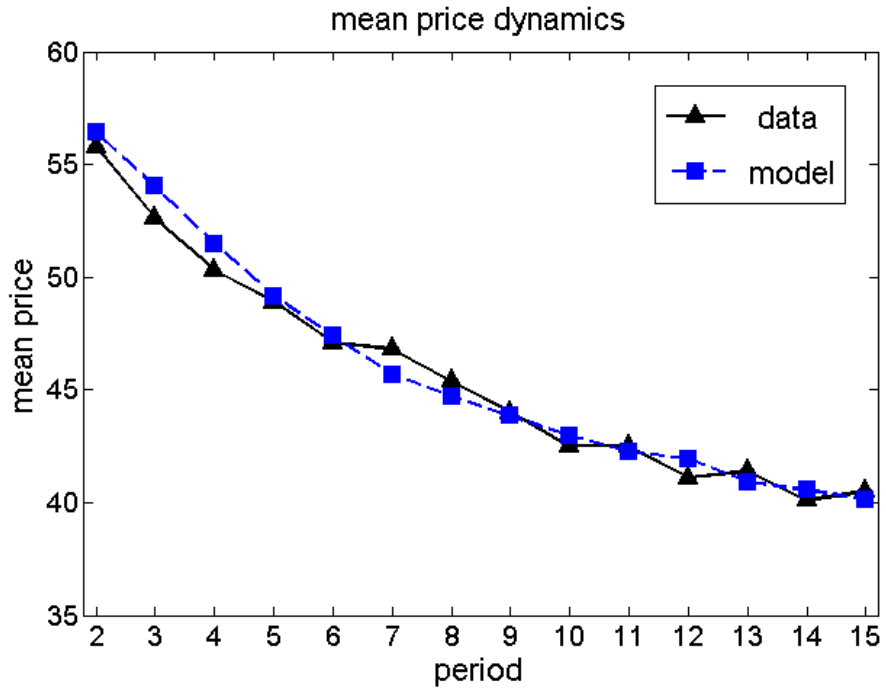
Figure 2.3.1: Average price dynamics: data observation and GPAL prediction
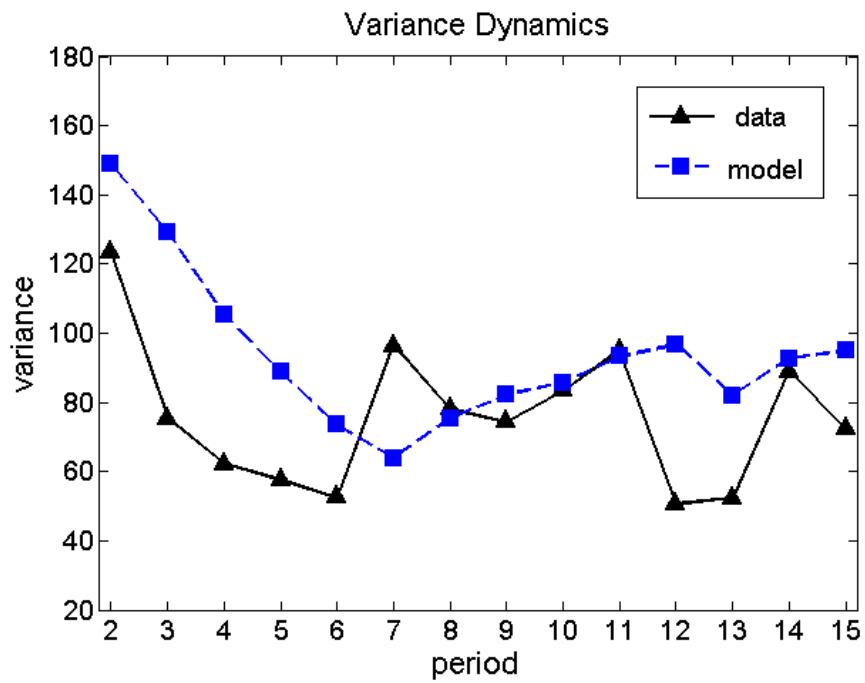
Figure 2.3.2: Price variance dynamics: data observation and GPAL prediction
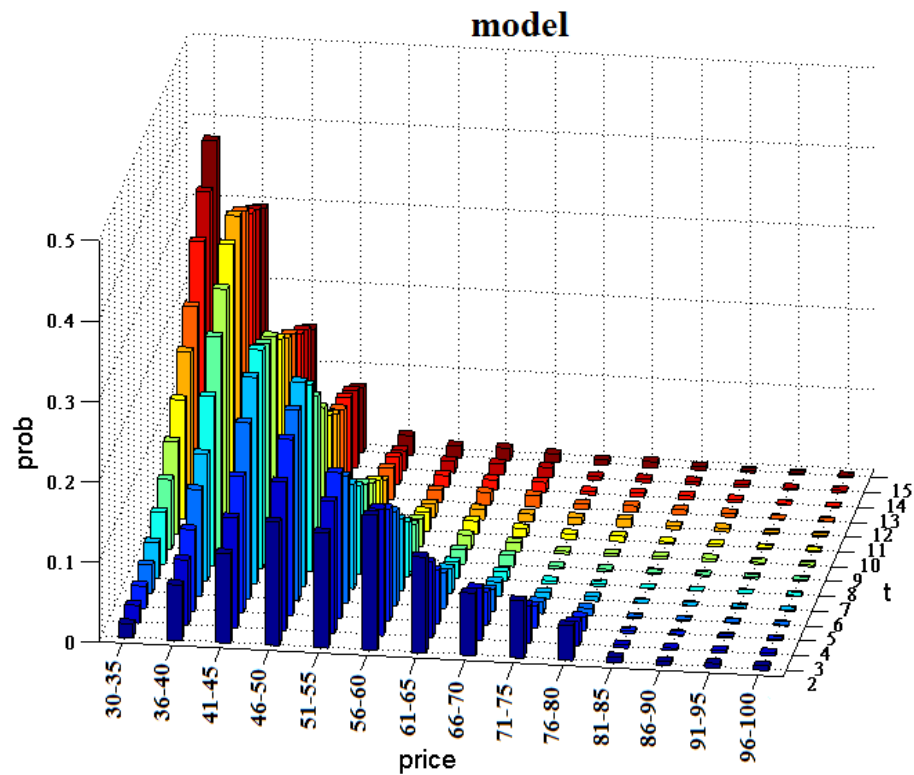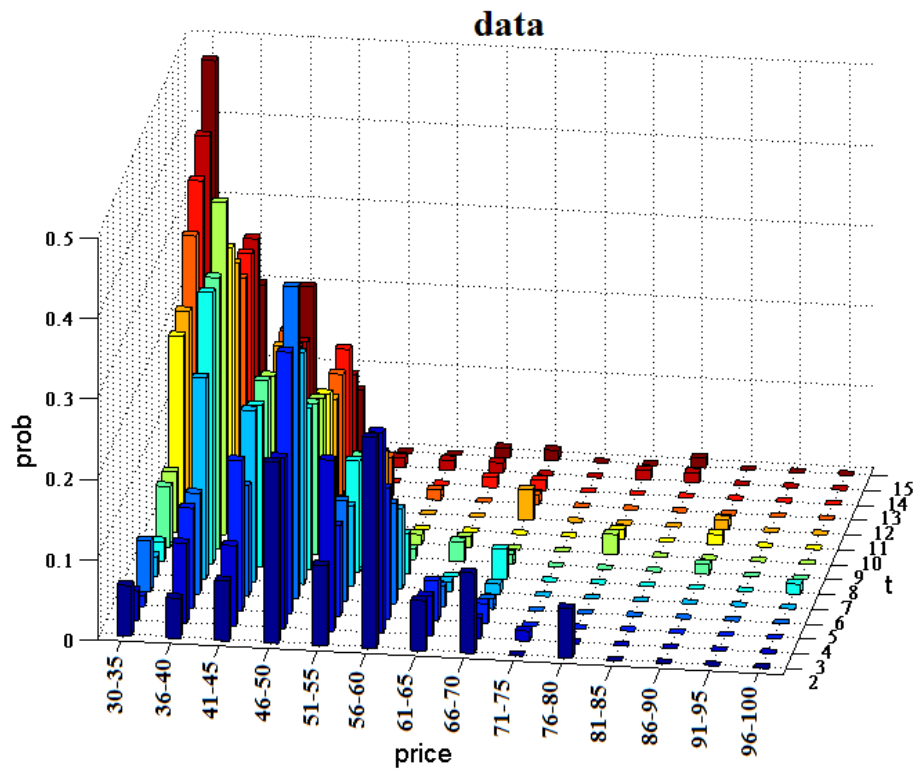
Figure 2.3.3: Distributions of prices - Data and Model

29

2.4   Conclusion

In this Chapter we proposed a simple extension to the payoff assessment learning model of Sarin & Vahid (1999) to explain sellers' learning behaviour in homogeneous-product Bertrand duopoly experiments with perfect information. The most prominent feature of the model is that it stresses the importance of the forgone payoffs for unchosen actions. In a broad sense, there are two categories of forgone payoffs. The first category is the forgone expected payoff, which is associated with games where the players are randomly matched in each period. Although being adaptive this component could also been seen as pointing to the future. Someone might do the following calculation. What profit will I get on average if I post this price next period and the others do not change their behaviour? The second category is the forgone payoffs relating to a player's own experience. In this case, the forgone payoff for choosing a price is the payoff a seller could have received if she had chosen that price and played with the same opponent. By combining both types of forgone payoffs, our model assumes that sellers use all relevant information and enables us to estimate the relative impact of different types of information on learning. For the Bertrand duopoly experiments we study, it seems that the information of a seller's actual opponents' choices overwhelmingly dominates the information of the potential competitors' prices in driving a sellers' price adjustments. However, the small but significant impact of the belief-learning factor considerably improves the fit of the model.

Chapter 3

Number of Sellers and Quantal Response Equilibrium Prices

3.1 Introduction

In homogeneous-product Bertrand oligopoly markets with identical sellers and perfectly informed buyers, Nash equilibrium asserts that all sellers set prices uniformly at the marginal cost and the prices are independent of the number of sellers in the market. Paradoxically, empirical observations provide overwhelming evidence against the Nash equilibrium prediction (e.g., Baye et al. 2004; Clay et al. 2001; and Lach 2002). Informational frictions on the side of the buyers and heterogeneity among sellers are intensely used in models to rationalize empirically observed price dispersion as a market equilibrium (e.g., Burdett & Judd 1983; Reinganum 1979; and Varian 1980). However, persistent price dispersion is still prevalent in experimental Bertrand markets where the above factors are deliberately controlled for (e.g., Abrams et al. 2001; Bayer & Ke 2011; and Dufwenberg & Gneezy 2000), as well as in online-shopping markets where search and information costs are negligible (e.g., Baye & Morgan 2004). Several recent studies have shown that the concept of Quantal Response Equilibrium of McKelvey &

Palfrey (1995) can effectively explain the price dispersion observed in laboratory markets (e.g., Baye & Morgan 2004; Capra et al. 2002; and Dufwenberg et al. 2007). In a Quantal Response Equilibrium (QRE), sellers are assumed to be boundedly rational and play noisy best responses to their beliefs about other players' strategies. Moreover, the beliefs held by all sellers are correct and prices with higher expected payoffs are more likely to be played. This Chapter investigates how the changes in the number of sellers affects QRE market prices.

We study two different but closely related specifications of QRE, namely, the power-function specification and the logistic specification. In the literature, the two specifications are frequently used and considered as very similar substitutes that lead to roughly equivalent approaches. For the power function specification, the probability of posting a price is proportional to a power function of its expected payoff.[1] In contrast, for the more widely used logistic specification the choice probabilities of prices are proportional to an exponential function of the corresponding expected payoffs (see McFadden 1973). Since closed form solutions are not readily available for both specifications we use a simple Cournot adjustment algorithm, which by design yields a QRE if it converges. The Cournot dynamics can not only be used to iterated towards an equilibrium but can also be interpreted as a learning rule towards equilibrium. We simulate quantal-response dynamics in Bertrand Oligopolies with different model specifications and number of firms. We find that, quite surprisingly, the two specifications result in opposite effects of an increase in the number of sellers on the QRE market-price distributions. Under the power-function specification, an increase in the number of competing sellers results in a decrease in average market price. In contrast, under the logistic specification, having more sellers can increase the average market price.

The remainder of the Chapter is structured as follows. The next section describes the theory. It lays out in detail the two Quantal Response Equilibrium approaches

---

[1]The power-function specification is a generalized form of the classic Luce (1959) discrete model in which the power is equal to one.

of modeling price dispersion in Bertrand oligopoly markets. Section 3.3 presents the simulations of the QRE models. Section 3.4 concludes.

3.2   Theory

We consider a Bertrand oligopoly market with n sellers engaged in price competition for selling exactly one unit of a homogeneous product. All sellers produce at identical and constant marginal costs c which, without loss of generality, is normalized to zero. Let v be the choke-off price above which demand is zero. Sellers choose prices simultaneously and independently from the non-trivial price interval $P \equiv [0, v]$. A seller whose price is lower than all other sellers' prices serves the market alone and earns the corresponding profit. If a seller is not among the lowest priced sellers, she does not sell and earns zero profit. In the case that m sellers are tied at the lowest price, they split the market equally. Given the price profile $(p_1, p_2 \cdots, p_n)$, for seller $i \in N \equiv \{1, 2, \cdots, n\}$ the payoff function can be formalized by

$$\Pi_i(p_{1,} p_{2,} \cdots, p_n) = \begin{cases} p_i & \text{if } p_i < p_j \forall j \neq i \\ \frac{p_i}{m} & \text{if } m \text{ sellers are tied at the lowest pirce } p_i \\ 0 & \text{otherwise} \end{cases}$$

The only symmetric Bertrand-Nash equilibrium for the above oligopoly game is for all sellers to set prices at zero. The Bertrand-Nash equilibrium prediction is based on the assumption that sellers are perfectly rational and behave optimally. However, it is likely that the players' choices are prone to errors and exhibit noise due to inexperience, computational limits, or mood shocks. This implies that sellers may not be able to play best responses with certainty. Quantal-response choice rules extend the Nash

equilibrium logic by allowing sellers to be boundedly rational and to choose suboptimal strategies with positive probability. Let the strategy of seller $i \in N$ be a probability measure over the price set P, as $F_i : P \longrightarrow [0, 1]$. Further let $B_i(F_{-i})$ be seller i's belief about the probability measure of her opponents' strategies. For simplicity, we impose the symmetry assumption so that all sellers adopt identical strategies. Accordingly, for seller i the expected payoff of choosing price p is expressed as

$$E\pi_i (p, B(F_{-i})) = p \left[1 - B_i(F_{-i}(p)) \right]^{n-1}. \tag{3.2.1}$$

Given the expected payoff function, the cumulative probability of seller i choosing price p, under the logistic specification, is stated as

$$F_i^{LS} (p, B_i(F_{-i})) = \frac{\int_0^p e^{\lambda \cdot E\pi_i(q, B(F_{-i}))} dq}{\int_0^v e^{\lambda \cdot E\pi_i(k, B(F_{-i}))} dk}, \forall p \in P, \forall i \in N. \tag{3.2.2}$$

Replacing $E\pi_i (p, B(F_{-i}))$ in Equation 3.2.2 by $\ln [E\pi_i (q, B(F_{-i}))]$ yields the corresponding power function specification as

$$F_i^{PS} (p, B_i(F_{-i})) = \frac{\int_0^p E\pi_i (q, B(F_{-i}))^\lambda dq}{\int_0^v E\pi_i (k, B(F_{-i}))^\lambda dk}, \forall p \in P, \forall i \in N. \tag{3.2.3}$$

Common to both specifications, the probability of a seller posting a price is positively related to the expected profit it yields. Better choices are played with higher probabilities but the best choice is not played with certainty. The parameter $\lambda \in [0, \infty)$ is used to measure the degree of bounded rationality in sellers' decisions. With $\lambda = 0$, sellers are completely confused and their decisions are insensitive to expected profit differences. In this case, a seller determines the price by drawing from a uniform distribution over the whole price set. As λ increases, sellers become more precise in making choices, and as λ tends to infinity they become fully rational and behave optimally. At a QRE, all sellers' strategies are quantal responses to their beliefs about the competing

34

sellers strategies and the beliefs are correct.

Definition. A strategy profile $F^{LQRE} \in [0,1]^n$ comprises a Logistic Quantal Response Equilibrium *(LQRE)*, if for all $i \in N$ and $p \in \mathrm{P}$, we have $F_i^{LQRE}(p) = F_i^{LS}(p_i, B_i(F_{-i}))$ and $B_i(F_{-i}(p)) = F_{-i}^{LQRE}(p)$. Similarly, a strategy profile $F^{PQRE} \in [0,1]^n$ comprises a Power-function Quantal Response Equilibrium *(PQRE)*, if for all $i \in N$ and $p \in \mathrm{P}$, we have $F_i^{PQRE}(p) = F_i^{PS}(p_i, B_i(F_{-i}))$ and $B_i(F_{-i}(p)) = F_{-i}^{PQRE}(p)$.

For the power function specification, Baye & Morgan (2004) show that a closed-form representations of symmetric F^{PQRE} can be obtained and comparative statics of the Quantal Response Equilibrium strategies can be analytically studied. [2]

Proposition 3.1. *(Baye & Morgan 2004) For any $\lambda \in [0, \frac{1}{n-1})$, the following comprises a symmetric PQRE:*

$$F_i^{PQRE}(p) = 1 - \left[1 - \left(\frac{p}{v}\right)^{1+\lambda}\right]^{\frac{1}{1-(n-1)\lambda}} \quad \forall p \in \mathrm{P}, \forall i \in I \qquad (3.2.4)$$

Two features of Proposition 1 are noteworthy. First, for a given $\lambda \in [0, \frac{1}{n-1})$, $F_i^{PQRE}(p)$ is uniquely determined. When $\lambda = 0$, prices are uniformly distributed over the price set. As λ approaches $\frac{1}{n-1}$, $F_i^{PQRE}(p)$ tends to one for all $p \in \mathrm{P}$, which corresponds to the Bertrand-Nash equilibrium outcome. Note that Equation 3.2.4 is only giving a valid equilibrium density for $\lambda < \frac{1}{n-1}$, as it would yield negative cumulative densities for a greater λ. For $\lambda \in [0, \frac{1}{n-1})$, an increase in n leads to a higher $F_i^{PQRE}(p)$

---

[2]For a detailed analysis and an experimental test of the comparative statics properties of the PQRE, see Dufwenberg et al. (2007).

for all $p \in (0, v)$, and thus to a decrease in the PQRE price in terms of first-order stochastic dominance.

Unfortunately, we are not able to derive a closed-form solution for the LQRE probability function, and thus cannot study the comparative statics properties of LQRE analytically. For this reason, in the next Section we simulate the model with both specifications and investigate the effects of increasing seller numbers from the simulated results.

3.3 Simulations

In the simulations we consider a discretized version of the game with price set P $\equiv$ $\{0, 1, \cdots, 100\}$. We have a design with four different seller numbers ($n = 2, 3, 4, 5$) with $\lambda = 0.15$. We use $\lambda = 0.15$ because according to Proposition 1, with $\lambda = 0.15$ as we can be sure from Equation 3.2.4 that a symmetric QRE exists for up to five competitors. Rather than calculating the QRE probability distribution of prices directly, we adopt a simple Cournot adjustment algorithm which, as we shall show, can provide us with meaningful dynamics that converge to QRE for all cases of interest. The Cournot processes start with a uniform probability measure of prices $F_{t=0}(p) = \frac{p+1}{101} \forall p \in$ P. For each of the following periods, $F_{t=1,2,3,\cdots}$ is defined as a quantal response to F_{t-1}. Formally, the Cournot process for the power function specification is given as

$$F_t^{PS}(p) = \frac{\sum_{q=0}^{p} E\pi \left(q, F_{t-1}^{PS}\right)^{\lambda}}{\sum_{k=0}^{v} E\pi \left(k, F_{t-1}^{PS}\right)^{\lambda}}, \forall p \in \text{P}, \forall t = 1, 2, 3 \cdots ; \tag{3.3.1}$$

similarly, the process for the logistic specification is stated as

$$F_t^{LS}(p) = \frac{\sum_{q=0}^{p} e^{\lambda \cdot E\pi \left(q, F_{t-1}^{LS}\right)}}{\sum_{k=0}^{v} e^{\lambda \cdot E\pi \left(k, F_{t-1}^{LS}\right)}}, \forall p \in \text{P}, \forall t = 1, 2, 3 \cdots . \tag{3.3.2}$$

If the Cournot process converges to a steady state, then the strategy profile in which all sellers adopting the corresponding mixed strategies at the steady state is necessarily a symmetric Quantal Response Equilibrium. This is because, if the process arrived at a steady state, the following two conditions that a QRE requires are both satisfied. First, all sellers are playing quantal responses to the belief that other sellers will keep their strategies. Second, the beliefs are correct because at the steady state no one will change her behaviour. For each of the four cases, we simulate the Cournot process for a thousand periods. Let δ_t^M be the absolute value of the maximum inter-temporal probability difference between period t and $t - 1$, which can be expressed as

$$\delta_t^M = \max_{p \in P} |f_t(p) - f_{t-1}(p)|,$$

where $f_t(p)$ is the probability density of price p at period t. If a Cournot process converges, then there is a t^* such that $\delta_t^M = 0$ for all $t \geq t^*$. As can be seen from Figure 3.3.1, for both specifications the Cournot processes converge in all $n = 2, 3, 4, 5$ scenarios.

For the power function specification, the results of the simulations are consistent with implications that can be derived from Proposition 1. Figure 3.3.2(A) shows the average price dynamics of the Cournot processes. As n increases from two to five, at the equilibria the average market price declines from 49.2 to 33.6. Figure 3.3.2(B) shows the corresponding probability distributions at the steady states ($t = 100$). As n increases, in the PQRE increasingly more probability mass is shifted from the high price domain (50 to100) to the low price domain (0 to 50).

Now we turn to the simulations of the logistic specification. Surprisingly, increasing the number of sellers affects LQRE in the opposite direction compared to the effects observed under PQRE. As shown in Figure 3.3.3 (A), the average LQRE market price increases from 37.8 in the duopoly scenario to 42.5 in the pentaopoly scenario. Figure

37

(A) Power Function Specification

(B) Logistic Specification
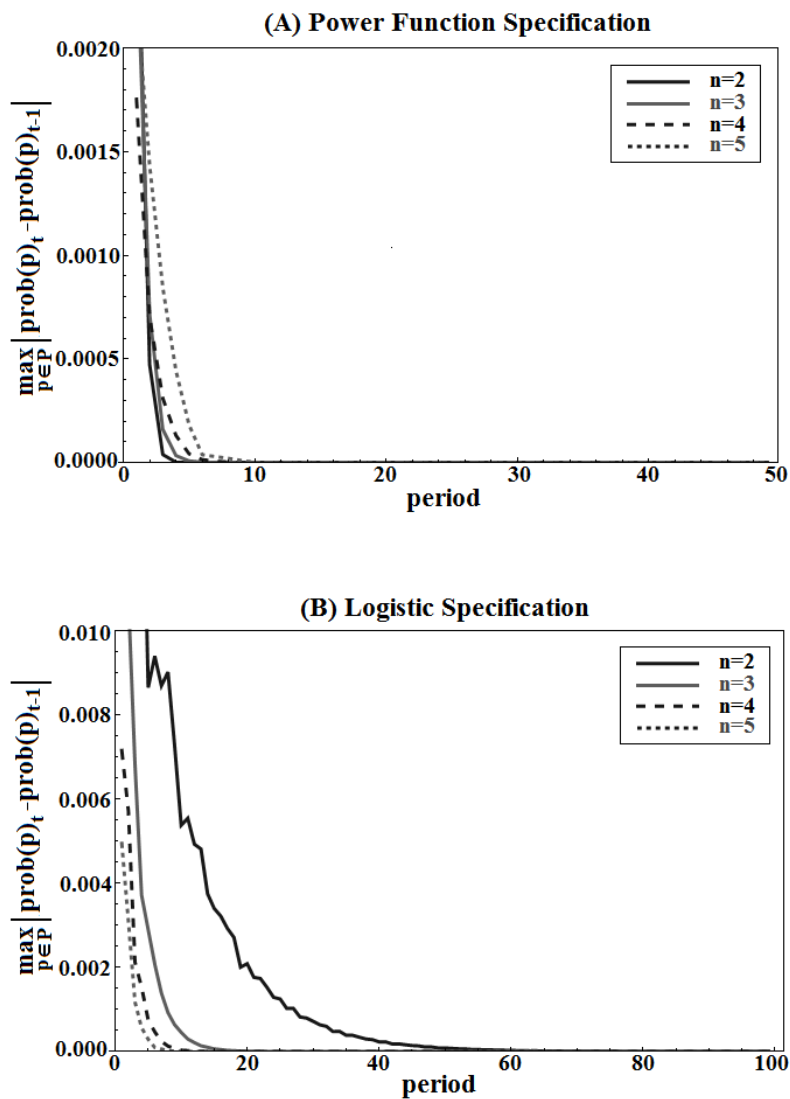
Figure 3.3.1: Time Series for the Maximum Absolute Value of Inter-temporal Changes in Price Densities with $(\lambda = 0.15)$

(A) Time Series for Average Market Prices

$\overline{p}(n=2) = 49.2$

$\overline{p}(n=3) = 44.6$

$\overline{p}(n=4) = 39.3$

$\overline{p}(n=5) = 33.6$

n=2
n=3
n=4
n=5

mean prices

period

(B) Probability Densities of Prices at t=100

n=2
n=3
n=4
n=5

prob(p)

price

Figure 3.3.2: Time Series of Average Market Prices (A) and PQRE Probability Densities of Prices (B) : Power Function Specification with $\lambda = 0.15$

3.3.3 (B) shows the corresponding probability distributions at the LQRE. This may help explain why increasing seller numbers lead to a less competitive market price with the logistic specification. In contrast to the PQRE, in LQRE as n increases, prices at the two ends of the price interval attract more probability mass, and those in the middle range are allocated lower mass. The increase in probabilities of the high prices is the main force that drives up the equilibrium average market price.
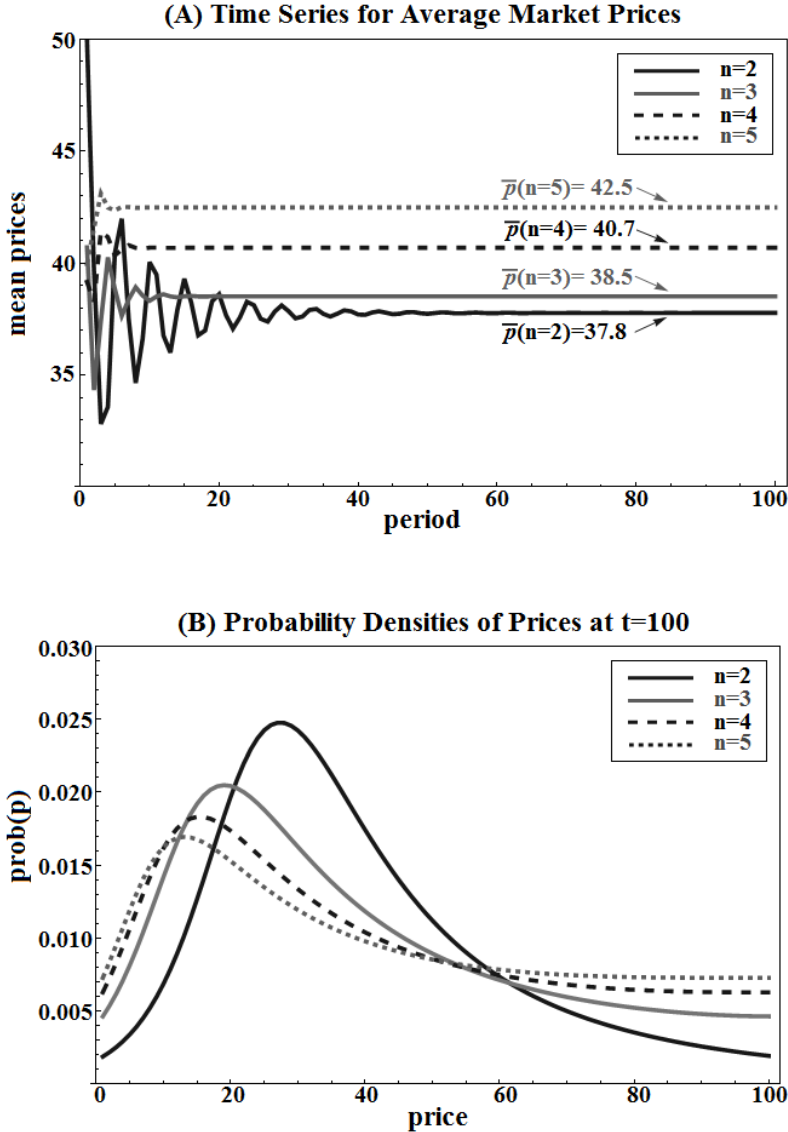
Figure 3.3.3: Time Series of Average Market Prices (A) and LQRE Probability Densities of Prices (B) : Logistic Specification with $\lambda = 0.15$

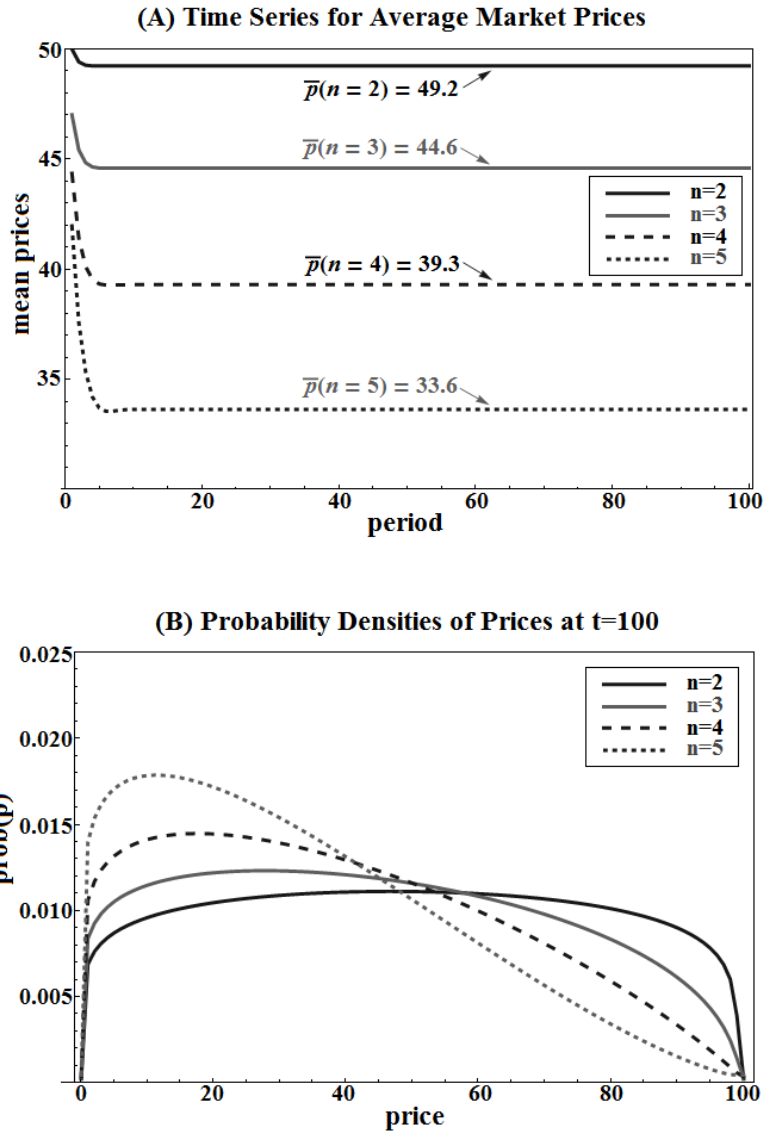So far we have focused on $\lambda < \frac{1}{n-1}$, where the condition of Proposition 1 is satisfied and the Cournot processes converge. It is also interesting to investigate how would a change in the number of sellers affect the Cournot dynamics if we relax the restriction and simulate the models with a higher λ. We conduct a new set of simulations using $\lambda = 0.4$ while keeping everything else unchanged. Figure 3.3.4 and Figure 3.3.5 present the results of simulations for the power function specification and the logistic specification, respectively. For the power function specification, when n equals to two and three (we still have $\lambda < \frac{1}{n-1}$ for these two cases), the Cournot processes still converge. However, when n equals three or four, the Cournot processes evolve cyclically after a few periods and the cycles persist over time.
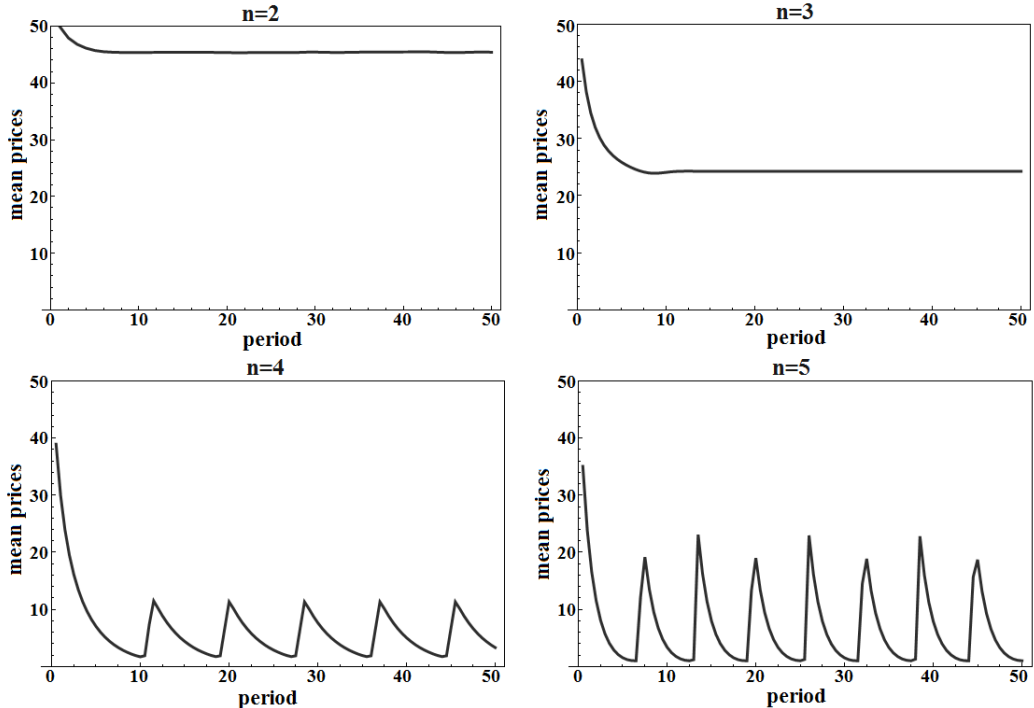
Figure 3.3.4: Time Series of Average Market Prices: Power Function Specification with $\lambda = 0.4$

41
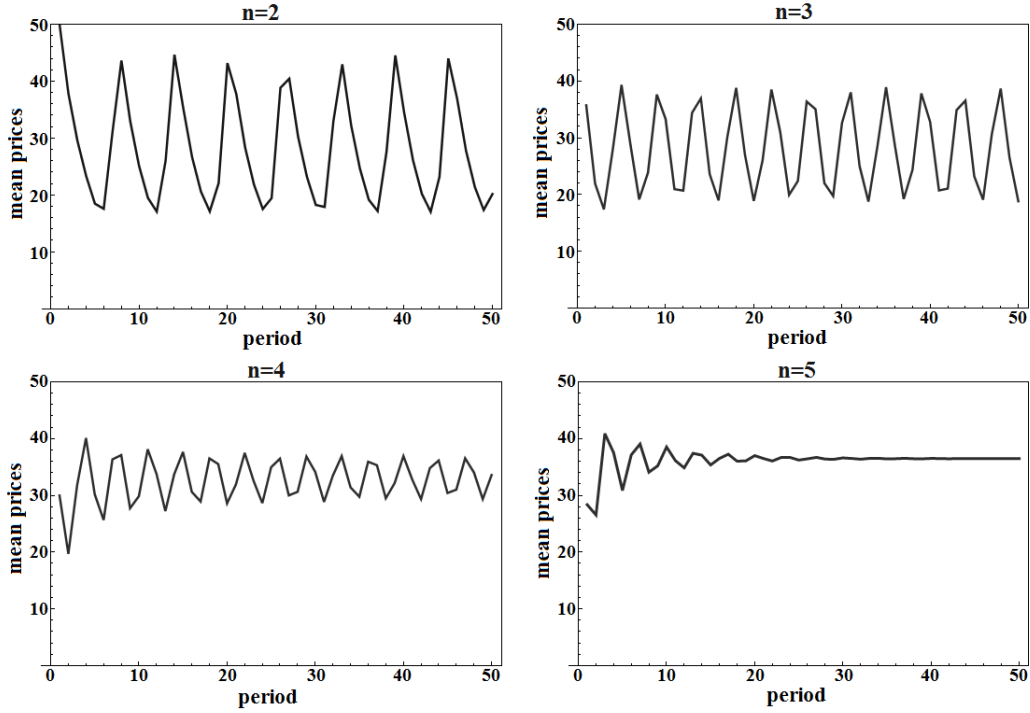
Figure 3.3.5: Time Series of Average Market Prices: Logistic Specification with $\lambda = 0.4$

For the logistic specification, the simulations still produce results contrasting to those obtained from the simulations with the power-function specification. In the duopoly, triopoly and quadropoly scenarios we observe persistent cycles of average market prices. However, increasing the number of sellers reduces the amplitudes of the price cycles and, when the number of sellers is increased to five the process converges after about 30 periods. We also ran simulations for markets with six to ten sellers, in all of these cases the Cournot process also converges, and the results indicate that as the number of sellers increases the speed of convergence increases. Again, increasing the number of competitors increases the market price. Therefore, we conclude that increasing the number of sellers may result in nonconvergence of the Cournot process under the power function specification, while can lead to convergence of the process under the logistic specification.

3.4 Conclusion

The conventional viewpoint that increased competition among sellers has the effect of reducing market prices has been challenged by many authors. Some models introduce search costs for consumers hunting for the lowest market price (e.g., Satterthwaite 1979; Stiglitz 1987; and Janssen & Moraga-González 2004). With more sellers competing in the market it is more costly for the consumers to succeed when searching. This effect can reduce search intensity, which in turn gives firms more market power. The possible result are increased prices. An alternative approach is to divide the consumers into two different types, for instance, loyal and swinging buyers in Rosenthal (1980) or, informed and uninformed buyers in Varian (1980). When facing intensified competition the sellers may have an incentive to exploit the loyal buyers and the uninformed buyers by charging a higher price. This paper demonstrates that the same phenomenon can arise in homogeneous product Bertrand oligopoly markets with identical sellers and perfectly informed buyers. The results of this Chapter also indicate that caution is necessary when choosing the quantal response specification to model Bertrand competition. The two dominant specifications which typically are seen as substitutes in modeling, lead to qualitatively vastly different results with respect to the impact of number of firms on price levels.

Chapter 4

Learning under limited information in laboratory Bertrand markets

4.1 Introduction

In experimental studies where people engaged repeatedly in some stage game, it is well known that the availability of information about other players' past actions can have substantial influence on an agent's strategic adjustment (e.g., Mookherjee & Sopher 1994; Huck et al. 2000; and Abbink et al. 2004). For a wide variety of games where players know only their own actions and payoffs, choice reinforcement learning (cf. Bush & Mosteller 1955; Erev & Roth 1998) and payoff assessment learning (cf. Sarin & Vahid 2001) models have been successfully used to explain observed behaviour.

Traditional choice-reinforcement learning (CRL hereafter) uses propensities to play a strategy. Initial propensities are given and lay outside the model. The choice probability for a certain strategy is determined by the propensity of this strategy relative to the propensity of the chosen strategy is updated according to a reinforcement function, while the propensity for all the unchosen strategies stays unchanged. In the payoff assessment learning model (PAL hereafter) subjects choose the strategy that they assess

subjectively to lead to the highest payoff. Again, initial assessments are assumed to be given and lie outside the model. PAL has in common with CRL that updating only occurs for the chosen strategy. So a player updates the payoff assessment for a chosen strategy according to an updating rule, weighs recent experience more strongly than experiences in the distant past. Both CRL and PAL models largely ignore an important factor that may have potent influence on learning: players' ex-post inferences about the payoff other strategies might have yielded. The omission only really matters in settings where players know the game and payoff structure. Note that knowing the game and payoff structure does not necessarily mean that subjects can infer the forgone payoff for every possible alternative strategy. In low-information environment such as ours, a player does not observe the strategy chosen by other players and might not be able to infer it from her own payoff. In such a situation we cannot use learning algorithms designed for high-information environments such as fictitious play. For this reason we will develop extensions to CRL and PAL that take into account what subjects can infer or might suspect about the forgone payoffs from unchosen strategies after observing their payoffs.

Our study is based on observations from a Bertrand duopoly experiment conducted by Bayer & Ke (2011). In the experiments, subjects were assigned to fixed roles as sellers and buyers, and randomly re-matched in each period to form market groups. In total subjects played fifteen periods. Over-all 186 subjects participated such that in each period 62 markets took place. A market consisted of a standard Bertrand price competition game. Two sellers simultaneous set integer prices between 30 (marginal cost) and 100 (reservation value for the buyer) to compete for a buyer wishes to buy one unit of the good in question. Buyers initially saw one of the two prices and could without incurring any cost uncover the other price with a simple mouse-click. Then buyers could either choose the seller to buy from or exit the market without buying.[1]

---

[1]In more than 99% of the cases the buyers sampled both prices and bought from the seller offering the lower price.

The payoff for successful seller was the difference between price set and cost, while an unsuccessful seller earned zero. The buyers' payoffs were calculated as their reservation value minus the price they paid if they bought from a seller, and were equal to zero if they did not buy. Before the experiments, participants were provided with detailed instructions. After each repetition, the only feedback revealed to the participants were their own payoffs which meant that sellers could infer only if they sold or not but not the exact price charged by the competitor.

The Bertrand-Nash equilibrium for the stage game requires sellers to set prices equal to 30 (marginal cost) or just above. The experimental results, however, differed markedly from the Bertrand-Nash prediction. The main features of the experimental results can be briefly summarized as follows: (1) Prices were persistently dispersed above marginal cost and the Bertrand-Nash prices were rarely observed; (2) The mean and median prices as well as the corresponding price variance decreased with repetition; (3) The sellers tended to increase or keep their prices unchanged after a successful sale and to reduce prices after failing to sell. Based on these observations, it is clear that the CRL or PAL models are not adequate. To illustrate, consider the case of a seller who has incurred a sales failure and obtained a zero profit. According to the assessment updating rules, for CRL there would be no change in the attractiveness for all alternatives, and for PAL the only change is that the payoff assessment for the chosen price is adjusted down. As a result, there would be little change in the seller's choice probabilities. This is obviously inconsistent with the experimental evidence.

Our model address the problem by extending the PAL model. We assume that the players update the assessments not only for the chosen prices using the realized payoffs, but also for all the unchosen prices for which the forgone payoffs can be inferred from the feedback received. More specifically, knowing their own prices, an unsuccessful seller can easily infer that all higher prices would also have led to failing to seller with a profit of zero. Likewise, a successful seller can infer that all lower prices would also

46

have led to a successful sale and hence the forgone payoffs for lower prices are simply the corresponding markups. Both successful and unsuccessful sellers can never know all forgone profits with certainty. A successful seller who does not know the price of the competitor cannot be sure what the profit for prices higher than that posted would have been, as she does not know if a sale would have happened. In the same way an unsuccessful buyer does not know the forgone profits for prices lower than the one posted. In order to make an assessment a player requires to have beliefs about the price set by the opponent conditional on the sale result. In what follows we assume a parametric form for these beliefs and estimate the parameters straight from the data. As we will show, the extended model organizes the data remarkably well at both aggregate level and individual level.

One learning model that bears significant relation to our approach is learning direction theory (cf. Selten & Stoecker 1986 and Selten & Buchta 1999). Directional learning is a qualitative approach that postulated that subjects ex-post rationally assess in which direction forgone payoffs would have been higher than the payoff actually realized. The theory points out that the realized payoff per se might not be the deciding factor in shaping how people adjust their actions. Instead, subjects adjust their behaviour according to their beliefs about how they can increase their payoffs. The theory therefore predicts that when the decision makers can infer from the feedback they received which actions could have led to higher payoffs, they are likely to adjust their behaviour in the direction of better performance. Strategic adjustments following this ex-post rationality principle have been shown to exist and persist in many experimental studies (see, e.g., Cason & Friedman 1997; Selten et al. 2005; and Bruttel 2009).

Grosskopf (2003) stressed the importance of combining choice reinforcement and directional learning when modeling learning behaviour both quantitatively and qualitatively is the aim. Our model goes a step beyond such combination: it embodies the idea that actions which a subject believes would have been better in the past are more

47

likely to be played, and more importantly, it allows us to estimate the degree of players' ex-post rationality based on the data.

The experience-weighted attraction learning (EWA) model of Camerer & Ho (1999) also emphasizes the impact of forgone payoffs of unchosen strategies on learning. EWA is a hybridized propensity updating model of learning which is suit for games with perfect information. It captures key features of choice-reinforcement learning and weighted fictitious learning and contains these models as special cases. In EWA, the forgone payoffs for unchosen strategies are weighted using a parameter $\delta \in [0, 1]$, while the realized payoff bears a weight of one. Moreover, a player's experience of playing the game is also incorporated in the attraction updating rule. Camerer & Ho (1999) show that, in terms of data fitting, EWA outperforms choice-reinforcement learning and weighted fictitious learning for a variety of experiments of games. For our limited-information setting, however, EWA is not directly applicable.

Our study is also closely related to the strategic similarity approach of Sarin & Vahid (2004).[2] The payoff assessment model allows subjects to update the assessments for the unchosen strategies that are similar to the chosen strategies. Specifically, Sarin & Vahid (2004) adopt the idea that similar actions expect similar payoffs and allow the update of assessments for strategies within a range of the chosen strategy. They use the Barlett and Parzen similarity function (e.g.,Brockwell & Davis 2009) to weigh the payoffs for unchosen strategies. According to the similarity functions, the closer a strategy is to the chosen strategy, the more weight the forgone payoff will have when the payoff assessment for this strategy is updated. While the strategic similarity approach is more appropriate for the situation where subjects are ignorant about the game environment, our approach is better suited for situations where players know the structure of the game but not the opponents' past strategy.

The Chapter is structured as follows. Section 2 summarizes the Bertrand duopoly

---

[2]See also Bayer et al. (2013) for its application in a voluntary contribution game and Chen & Khoroshilov (2003) for the use in a cost sharing game.

experiments of Bayer & Ke (2011). Section 3 presents the model. Section 4 shows the estimation of the model and compares the model predictions to the experimental data. Section 5 concludes with a discussion about the model's convergence properties in the long run.

4.2 Experiment

The experiment of Bayer & Ke (2011) was conducted in the Adelaide Laboratory for Experimental Economics (AdLab) at the University of Adelaide. Participants were mainly university students studying for undergraduate or postgraduate degrees in a variety of disciplines. In total 186 participants were recruited. Participants were asked to play fixed roles as sellers and buyers in a Bertrand duopoly market game for thirty periods. The purpose of Bayer & Ke (2011) was to investigate the effects of exogenous cost shocks on market price. So they conducted a two-phase experiment, with 15 periods for each phase. The only difference between the two phases was the sellers' marginal costs. In the first phase the marginal cost was 30 for all sellers and all treatments. After the first phase, in different treatments the marginal costs either went up to 50, was kept unchanged at 30, or declined to 10. Since the cost shock effects are beyond the interest of this paper, we focus on the first phase (15 periods) of the experiment where the cost shocks had not been imposed. Note that the cost shock was unanticipated such that it should not have any impact on play in the first periods. However, subjects knew that the experiment would run for 30 periods. At the beginning of each period, sellers and buyers were randomly matched to form markets. There were 62 markets and each market consisted of two sellers and a buyer. In each period, sellers with unit cost $c = 30$ simultaneously and independently post a price from the price set

$P \equiv \left\{ 30, 31, \cdots, 100 \right\}$ at which they offer to sell the product. Afterwards, the buyer with unit demand and reservation value $v = 100$ enters the market, learns one price and can click to see the other price without any cost. Then the buyers has to decide either to buy from one of the sellers or to leave without buying. After that a period ends. The payoffs for successful sellers are equal to their prices minus the cost, while the profits of unsuccessful sellers are zero. For a buyer, the payoff is v minus the price she paid if she bought and zero if she did not buy. After each period, the only information revealed to participants are their own payoffs. At the end of the experiments, the payoffs were aggregated over all periods and exchanged for real money at a fixed exchange rate. On average, the participants earned about twenty Australian Dollars for about one hour of their time.

Figure 4.2.1 shows the frequency distributions of the posted prices for all periods. Each bar represents the relative frequency for five consecutive prices to be charged as labeled at the ticks.[3] Prices are significantly dispersed above the marginal cost for all periods and the Bertrand-Nash equilibrium prices are rarely observed throughout. Further, the price distributions exhibit remarkable temporal variation, as there is an obvious tendency of prices steering away from the clusters above 60 toward those below 60.

---

[3]The lowest bin contains six prices as the number of strategies was not divided by five.
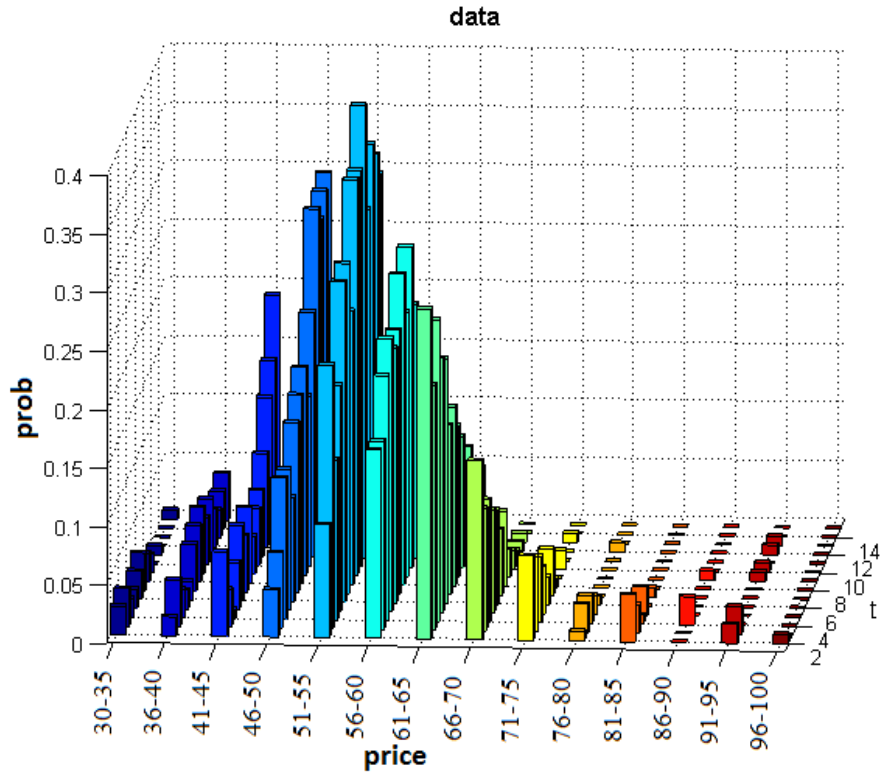
Figure 4.2.1: Price distributions by period

The tendency of price adjustment becomes clearer by plotting the dynamics of the mean and median market prices in panel (A) of Figure 4.2.2. Both mean and median prices start off at about 60 and decrease gradually with repetition. After nine periods, the speed of price reductions decreased markedly. Actually, the median price remains at 50 from period 11 to period 14. There is a similar trend for the variance of the prices (see plot (B) of Figure 4.2.2). The variance decreases quickly from 156 in period one to 77 in period six, and then slowly to 37 in the last period.

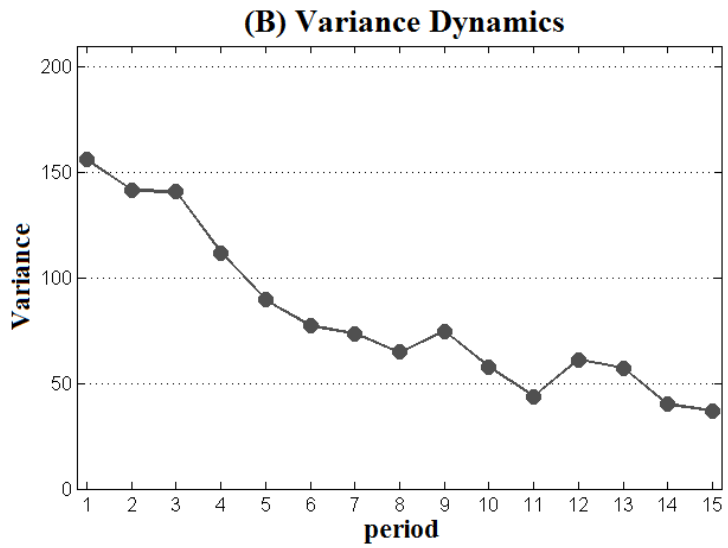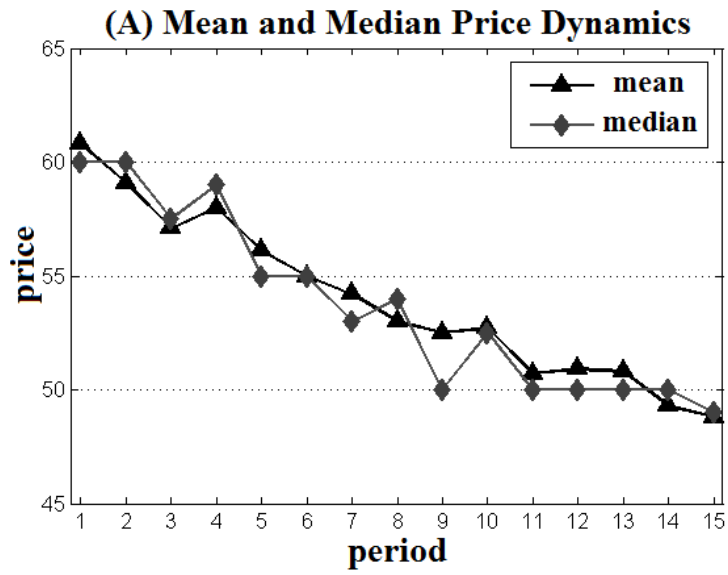Figure 4.2.2: Evolution of the mean and median prices (A), and the evolution of price variances (B)
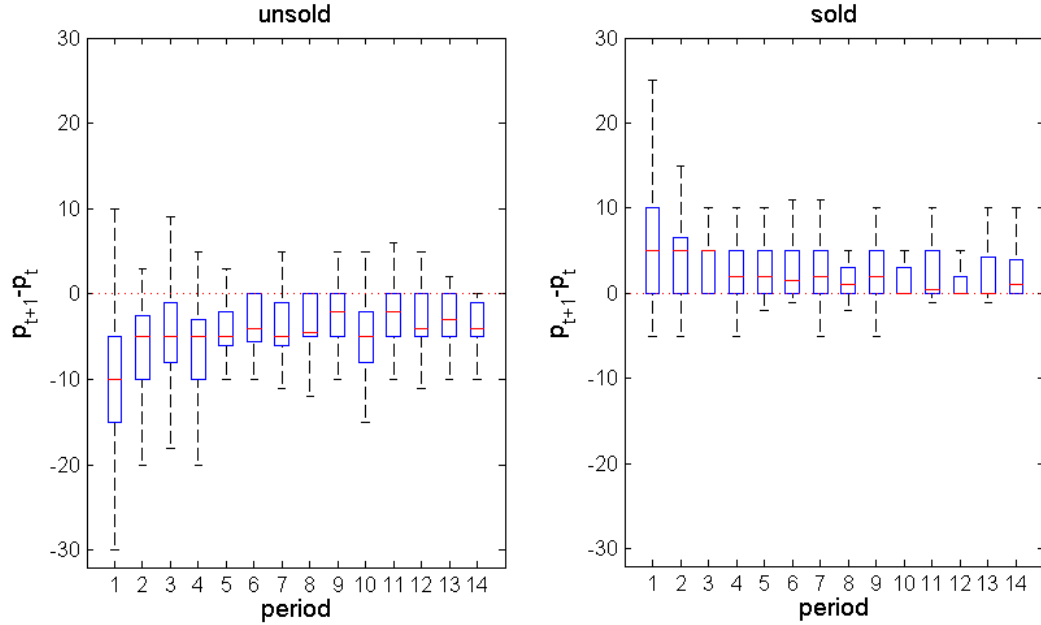
Figure 4.2.3: Box plots of price adjustments by sales outcomes(outliers excluded)

At the individual level, the most striking observation is that the sellers' price adjustments depend heavily on whether they sold their unit or not in the previous period. Figure 4.2.3 shows the interquartile box plots of the sellers' intertemporal price changes in all periods, respectively for the unsuccessful sellers (the left panel) and the successful sellers (the right panel). The red bars in the boxes depict the median price adjustments. As can be seen from the plots, the sellers, who were not able to sell in the period before, usually reduced their prices. Those sellers, who experienced a sale success typically adjust prices upwards or keep their prices unchanged. With other words, the intertemporal price adjustments observed present strong evidence for the sellers being ex-post rational in the sense of direction learning theory.

4.3 Model

In this section we propose a simple learning model base on the experimental observations. Our model differs from the payoff assessment tradition only in its incorporation of the forgone payoffs for unchosen strategies. For the chosen price, the payoff assessment for period $t + 1$ is defined as the weighted sum of the corresponding payoff assessment for period t and the realized payoff from period t. In contrast, for the unchosen prices, the assessments for period $t + 1$ are equal to the weighted sum of their respective assessments for period t and the forgone payoffs that a player believes these prices could have earned in period t. More specifically, for a successful seller, the forgone payoff for a lower price is its markup over the marginal cost because a lower price would certainly also have led to a sale if posted. The forgone payoff for a higher price, in contrast, is defined as the possible markup multiplied by the seller's believed probability that a sale would occur if this price were posted. Similarly, for an unsuccessful seller, the forgone payoff to a higher price is zero, and that to a lower price is its markup times the seller's believed sales probability associated with posting that price. Moreover, the sellers believed sales probabilities should satisfy the simple restriction that is decreases weakly with a higher price. Further to that, we would expect the probability of a sale to be zero if a price of v is charged. The probability of a sale for a price smaller than c should be one. Additionally, we would like to allow one degree of freedom for each for the reactivity of the beliefs for successful and unsuccessful sellers.

4.3.1 Learning rule

Let $A_t^i = (A_t^i(c), A_t^i(c+1), \cdots, A_t^i(v))$ denote the vector of payoff assessments of seller i in period t. The initial assessment for each $p \in \mathrm{P}$ is denoted by $A_{t=0}^i(p)$. After period t, the only feedback seller i obtains is her own payoff that resulted from the price

she posted. In other words, sellers know only whether or not the prices they posted have yielded a successful sale. Given the feedback from period t, seller i updates her assessment for price p by the following rule:

$$A_{t+1}^i(p) = \phi A_t^i(p) + (1 - \phi)\hat{\pi}_t^i(p); \ \forall i \in I, \ \forall p \in P. \tag{4.3.1}$$

In the equation, $\hat{\pi}_t^i(p)$ is the expected payoff that seller i believes she would have obtained if she had posted price p instead of p_t^i. Hence, after period t, the payoff assessment $A_{t+1}^i(p)$ is updated as the weighted sum of the previous payoff assessment $A_t^i(p)$ and $\hat{\pi}_t^i(p)$, where ϕ $(0 \leq \phi \leq 1)$ is the inertia parameter, which measures the relative weight the seller puts on her past payoff assessments.

For a seller who has enjoyed a sales success , $\hat{\pi}_t^i(p)$ is defined as

$$\hat{\pi}_t^i(p) = \begin{cases} p - c, & \text{if } p \leq p_t^i \\ (p - c)\left(\frac{v-p}{v-p_t^i}\right)^\alpha, & \text{if } p > p_t^i \end{cases}. \tag{4.3.2}$$

That is, if seller i succeeded at price p_t^i, then a lower price $p \leq p_t^i$ would also have yielded a sale, with a profit of $p - c$. A higher price, by contrast, would have resulted in a tradeoff between a higher markup and the risk of not selling. We use $(p - c)\left(\frac{v-p}{v-p_t^i}\right)^\alpha$ to capture such a tradeoff. This formulation satisfies the structure of the properties we require. The first part is the markup. The second part, $\left(\frac{v-p}{v-p_t^i}\right)^\alpha \in [0, 1]$, is used to capture the seller's (subjective) expected probability of sales success if she had posted a higher price p. Here $v - p_t^i$ is the length of the interval for prices that are greater than the posted price p_t^i, and $\frac{v-p}{v-p_t^i}$ is the proportion of prices that are above p. Accordingly, the closer p is to p_t^i, the higher is the probability the seller expects that he would have sold. The parameter $\alpha \in [0, \infty)$ is introduced to capture the sensitivity of a successful sellers' believes are towarding price increases. When $\alpha = 0$, the seller is extremely insensitive and believes that she would have sold at any higher price. On the other hand, when
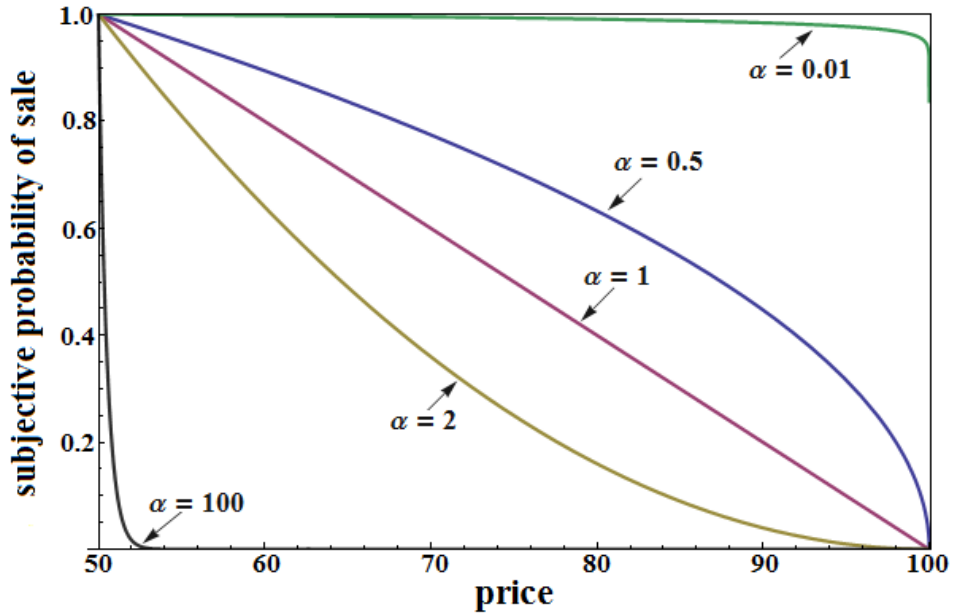
Figure 4.3.1: Subjective probability of selling at a higher price: an example with $p_t^i = 50$.

$\alpha \to \infty$, she is extremely sensitive and believes that any price increase would definitely have led to a failure to sell. Figure 4.3.1 provides an example of $\left(\frac{v-p}{v-p_t^i}\right)^\alpha$ with $p_t^i = 50$ for different values of α.

For sellers, who incurred a sales failure, $\hat{\pi}_t^i(p)$ is defined as

$$\hat{\pi}_t^i(p) = \begin{cases} 0, & \text{if } p \geq p_t^i \\ (p-c)\left(\frac{p_t^i-p}{p_t^i-c}\right)^\beta, & \text{if } p < p_t^i \end{cases}. \tag{4.3.3}$$

If a seller failed to sell at p_t^i, a higher price would also have led to a failure. In contrast, a lower price would have provided a chance to win the consumer. Similar to the case of successful sellers, $\left(\frac{p_t^i-p}{p_t^i-c}\right)^\beta \in [0,1]$ is used to capture an unsuccessful seller's (subjective) expected probability of sales success if she had posted a lower price. The sensitivity parameter $\beta \in [0,\infty)$ serves the same purpose as α does for the case of successful sellers. The closer is p to p_t^i, i.e. the higher price p is, the lower is the expected likelihood of having been able to sell at p .

56

Given the above definition of $\hat{\pi}_t^i(p)$, for both the successful and unsuccessful sellers, a smaller α or β implies larger expected forgone profits of $\hat{\pi}_t^i(p)$ for prices in the "could-have-been-better" direction compared to the other prices. In other words, we are more likely to observe successful sellers increasing their prices and the unsuccessful sellers reducing their prices if the estimates of α and β are small than when they are big. To see this, consider the extreme condition with α and β equal to zero. For the successful sellers $\hat{\pi}_t^i(p)$ will be equal to $p - c$ for all prices and thus higher prices will have a higher increase in their payoff assessment. This is because higher prices lead to higher margins. For unsuccessful sellers, for $p \geq p_t^i$ the payoff assessment will not be changed, but for $p < p_t^i$, the change in payoff assessment will be $\hat{\pi}_t^i(p) = p - c$, which is maximal. Increasing values dampen the directional learning effect, since then the changes in payoff assessments become smaller.

4.3.2 Choice rule

If a seller's price choices are not affected by random factors such as mood shocks or noise, then she would always choose the strategies that she assesses to have the highest payoffs. Now suppose that at period t, the seller experiences identically and independently distributed mood shocks $\varepsilon_t = (\varepsilon_t(c), \cdots, \varepsilon_t(p))$. Denote the shock-distorted payoff assessments as $\tilde{A}_t^i(p) \equiv A_t^i(p) + \varepsilon_t^i(p), \forall p$. At period t, instead of selecting the price that maximizes $A_t^i(p)$, we assume that the seller selects a price p if

$$\tilde{A}_t^i(p) > \tilde{A}_t^i(m), \forall m \neq p, \forall m \in \mathrm{P}.$$

We use the logit probabilistic choice model to account for the above feature of price choice. With this we implicitly assume that each $\varepsilon_t^i(p)$ follows a Gumbel or type I

extreme value distribution.[4] With this specification, at period t, the probability of seller i choosing price p is

$$f_t^i(p) = \frac{e^{\mu_t \cdot A_{t-1}^i(p)}}{\sum_{k=c}^v e^{\mu \cdot A_{t-1}^i(k)}}$$

where

$$\mu_t \equiv \lambda + (t-1)\gamma. \tag{4.3.4}$$

Parameter μ_t is used to measure how sensitive the sellers are to the differences in payoff assessments of different prices, or in the mood shock formulations how strong the moods shocks are. When μ_t is positive, the logit model has an intuitively appealing property for modeling decisions: the strategies that being assessed to have higher payoffs are more likely to be selected. As μ_t approaches infinity, the firms will always choose the prices with the highest payoff assessment. On the other hand, when μ_t equals to zero, a seller becomes fully ignorant or confused about the game she is playing and makes random choices. Differing from most previous applications of the logit model where parameter μ_t was taken as a constant across all periods, we endogenize μ_t as 4.3.4 for the following reasons.[5] First, while playing the game repeatedly the players may accumulate experience so they are likely to make more precise choices over time. It is common in experimental applications that the period-specific estimations of μ_t often yield higher values for later periods than for earlier ones (see McKelvey & Palfrey 1995 and Dufwenberg et al. 2007). Second, allowing μ_t to vary over time may help us to disentangle two different learning effects: the learning effect of accumulating experience and making more precise decisions, and the effect of the payoff assessment learning following Equation 4.3.1.

---

[4]With the Gumbel extreme value distribution (see Akiva & Lerman 1985), the density function for each $\varepsilon_t^i(p)$ is $f(\varepsilon_t^i(p)) = \exp[-\varepsilon_t^i(p) - \exp(-\varepsilon_t^i(p))]$ and the associating cumulative distribution is $F(\varepsilon_t^i(p)) = \exp[-\exp(-\varepsilon_t^i(p))]$.

[5]We also tried a nonlinear version using a harmonic function such as $\mu_t \equiv \lambda + \gamma \sum_1^t \frac{1}{t}$, which assumes that a player's sensitivity toward the payoff assessment differences is an increasing/decreasing function of t at a diminishing rate. It makes little difference in the estimates.

4.4 Estimation

4.4.1 Initial payoff assessments

The only remaining task for us before estimating the model parameters is to pin down for all players $i \in I$ the initial payoff assessment vector $A^i_{t=0}$. For games with small pure strategy sets $A^i_{t=0}$ could be estimated directly with the model (e.g., Camerer & Ho (1999); Erev & Roth 1998). With large pure strategy sets, however, estimating the initial assessments introduces massive computational complexity as well as a large number of degrees of freedom, which will almost certainly lead to over-fitting. A simple alternative (e.g., Chen & Tang 1998; Chen & Khoroshilov 2003) widely used is to assign to all pure strategies an uniform initial assessment that is equal to the average payoff earned by all players in period one. However, the null hypothesis of uniform initial assessments being compatible with the first period behaviour is rejected by a Kolmogorov–Smirnov(K-S) test at the 5% significance level (p-value= 0.003). Ho et al. (2007) adopt the idea of cognitive hierarchy theory (Camerer et al. 2004) where players are categorized as step k thinkers: step 0 players randomize, step 1 players best respond to step 0, and step k players best responds to their beliefs of the distribution (usually assumed to be a Poisson distribution) of step 0 to step $k-1$ thinkers. As an average of step 1.5 has been tested to fit well for a wide range of experimental data, Ho et al. (2007) set the initial payoff assessment of a strategy as its expected payoff while being played against step 1.5 thinkers. In our context, however, this approach is also not ideal. Best responses played against a uniform distribution (step 1) leads to a payoff distribution that is nice and smooth, and that can be easily derived.[6] Nevertheless, for step 2 players, in the case of playing against step 1 players, the expected payoffs for different prices change in an undesirable discrete manner: prices less than 65 (the

---

[6]However, the assumption that all sellers are step1 thinkers is rejected by K-S test at 10% significance level (p-value= 0.072)

price that step 1 players choose) yield expected payoffs equal to the markups and prices above 65 yield zero expected payoffs.

We use a reverse engineering approach in setting $A^i_{t=0}$. Specifically, we trace back for an appropriate price distribution to which the best response is the mode price observed in the first period data. It turns out that Poisson distribution serves our purpose quite well. The Poisson distribution has an appealing property of having only one parameter κ: let the markup of a price be $m \equiv p - c \in \{0, 1, \cdots, v - c\}$, then the probability function of the Poisson distribution with mean κ is given by

$$f(m) = \frac{\kappa^m e^{-\kappa}}{m!}, \forall m.$$

For the first period, the mode price of the data was 60 (so, $p - c = 30$) and it is a best response to a Poisson price distribution with $\kappa = 37$ (so the mean price of the distribution is $c + \kappa = 30 + 37 = 67$). Accordingly, the initial payoff assessment for price $p = m + c$ is defined as its expected payoff for being played against the Poisson distribution with $\kappa = 67$:

$$A^i_{t=0}(p) = (p - c) \left[1 - \sum_{q=0}^{p-c} \frac{\kappa^q e^{-\kappa}}{q!} \right], \forall p \in \mathrm{P}.$$

This approach is validated by the K-S test (p-value= 0.452) and fits the first period data way better than all the above mentioned approaches. For simplicity, in the estimation we shall make the assumption that the initial payoff assessment vectors are the same across all players.

4.4.2 Maximum likelihood estimation

For the estimation we search for the values of the parameters that maximize the log-likelihood of observing the experimental data. Formally, we cast about for the values of ϕ, α, β , λ and γ that maximize

$$\log(L) = \log\left[\prod_{t=1}^{15}\prod_{n=1}^{N} f_{i,t}(p)\right] = \sum_{t=1}^{15}\sum_{n=1}^{N}\log\left[\frac{\exp\left[\mu_t \cdot A_{t-1}^i(p)\right]}{\sum_{k=c}^{v}\exp\left[\mu_t \cdot A_{t-1}^i(k)\right]}\right]$$

where N is the total number of sellers in the experiments. In order to examine the robustness of the estimates, we conduct estimations not only for the whole data set ($N = 124$, nine sessions[7]), but also for two subsets with 40 sellers (three sessions) and 84 sellers (six sessions), respectively.

	ϕ	α	β	λ	γ	$\log(L)$
Whole data set	0.838	1.303	0.236	0.079	0.034	-5674
(124 sellers)	(0.051)	(0.063)	(0.027)	(0.016)	(0.011)	
Subset I	0.847	1.254	0.333	0.063	0.041	-3784
(84 sellers)	(0.047)	(0.056)	(0.033)	(0.022)	(0.020)	
Subset II	0.813	1.392	0.129	0.093	0.026	-1873
(40 sellers)	(0.066)	(0.080)	(0.042)	(0.041)	(0.037)	

Table 4.4.1: Maximum likelihood estimates (standard errors in parenthesis)

Table 4.4.1 reports the maximum likelihood estimates. The numbers in the parenthesis are the standard errors obtained by numerical differentiation. The table also includes the values of log likelihood function, $\log(L)$, at the estimated parameters. As can be seen, the estimates are close across different data sets for all parameters and, the values derived from the whole data set are roughly the average of those from the two truncated data sets. The discount parameter $\hat{\phi}$ is about 0.8 which implies that the new experience attracts a significant weight and thus affect the price decision substantially. The sensitivity estimates $\hat{\alpha} = 1.303$ and $\hat{\beta} = 0.236$ turn out to be small enough to ac-

---

[7]There were between 12 and 18 sellers in a session.

count for the directional learning price adjustments. For unsuccessful sellers, based on the model, downward adjustment of prices is obvious because only the lower prices may attract positive forgone payoffs $\hat{\pi}_t^i(p)$. For successful sellers, the estimate $\hat{\alpha} = 1.303$ implies, at an aggregate level, a trend of upward adjustment of the prices, which is less strong than the downward adjustment in the other case. To see this, consider an example in which seller i has successfully sold at $p_t^i = 60$. In this case, with $\hat{\alpha} = 1.303$ the sum of $\hat{\pi}_t^i(p)$ for prices higher than 60 is $\sum_{p=61}^{100} \hat{\pi}_t^i(p) = 746.4$, which is much higher than the sum of $\hat{\pi}_t^i(p)$ for the lower prices ($\sum_{p=30}^{59} \hat{\pi}_t^i(p) = 435$). This indicates that for seller i the impluse to increase the price is higher than the impulse to reduce the price. This is the case for all winning prices that are below 70. Since we have rarely observed a winning price greater than 70 in the experiments, we can safely say that the parameter $\hat{\alpha} = 1.303$ conforms to the observed directional learning price adjustment observed for successful sellers. With $\hat{\beta} = 0.236$ the adjustment pressure is greater for unsuccessful sellers. This is consistent with reference dependent preferences that weighs losses more strongly than gains.
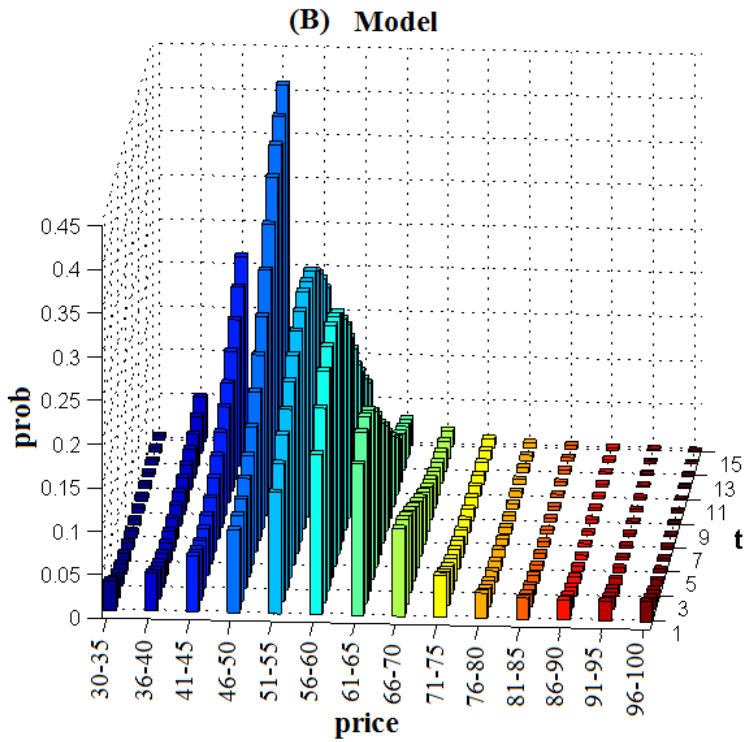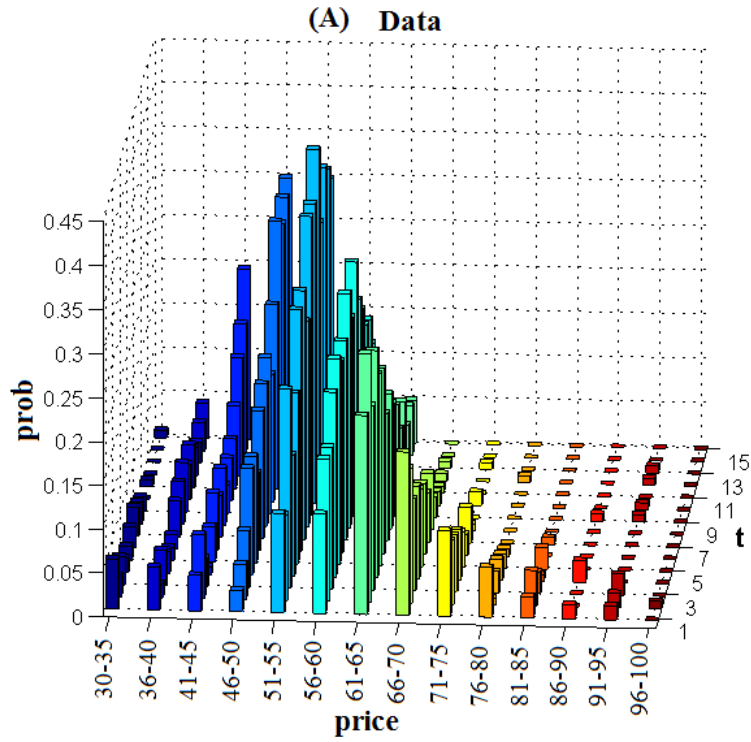
Figure 4.4.1: Price Dispersion: Data observation and Model Prediction

Figure 4.4.1 presents a comparison between predicted price frequencies of the model

63

(panel B) and the corresponding empirical frequencies (panel A). Overall, the model organizes the data remarkably well. The model precisely predicts the tendencies of relative frequency changes for most of the price clusters over time. The largest prediction error of the model is that it overpredicts the frequencies of cluster 46 to 50 and underpredicts those of cluster 51 to 55, especially for later periods.
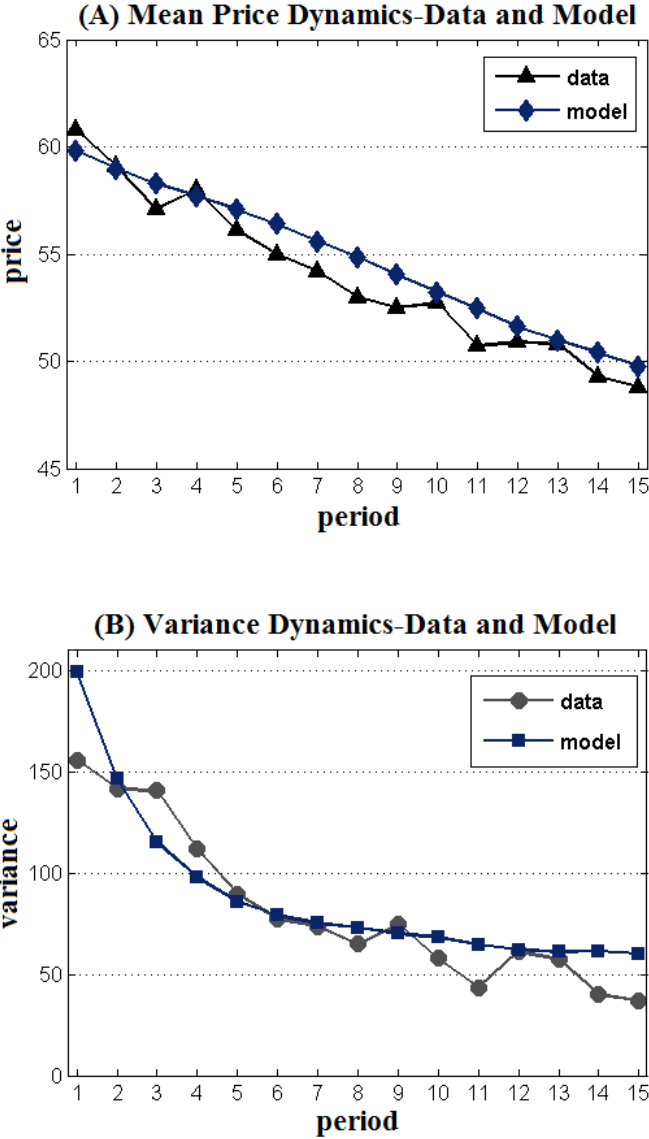

Figure 4.4.2: Mean (A) and Variance (B) dynamics: Data Observation and Model Prediction

Panel (A) of Figure 4.4.2 shows the time series of the mean price estimated by

the model, together with the real dynamics of average prices. Panel (B) shows the corresponding dynamics of price variances. Overall, we can see that the model does rather well in tracking both the means and variances of the prices. The only problem is that the model slightly overpredicts the mean prices for periods six to fifteen. We ascribe the errors to the model's overprediction of the frequencies for prices above 75. As showed in Table 4.4.2, for period six to period fifteen, the model predictions of the total frequencies for prices above 75 are systematically higher than the actual frequencies observed in the data (see also Figure 4.4.1).

period	6	7	8	9	10	11	12	13	14	15
$\Pr(p > 75)$ Model	0.027	0.023	0.021	0.019	0.018	0.017	0.016	0.016	0.016	0.016
$\Pr(p > 75)$ Data	0.008	0.008	0	0.016	0.008	0	0.016	0.08	0	0

Table 4.4.2: Time series of total probabilities for prices above 75

CRL and PAL models have advantages over the belief-learning models for allowing players' decisions to depend directly on their own past choices and payoffs. Our model enables us to predict players' idiosyncratic strategies for all but the first periods and evaluate the model's goodness of fit at an individual level.[8] To assess how well the model is tracking the individual's price dynamics, we compare for all sellers the prices they actually posted and the mode prices predicted by the model. There are 1760 ($N \times T$) pairs of posted prices and corresponding mode predictions. We then take the differences of each pair (mode prediction minus observed price) and plot the histogram of the frequencies for the pooled differences. Figure 4.4.3 shows the results. Overall, the mode predictions of the model fit the posted prices excellently. The mean of the differences is 0.52 and the median is zero. For about 41% of the cases, the predicted modes are within 2 units of the actually posted prices. Therefore, we feel confident

---

[8]Wilcox (2006) shows that for empirical estimations, if the comparison between reinforcement learning models and belief learning models are compared based only on the goodness of fit, the results will be biased in favor of the reinforcement learning models. This is because the reinforcement learning models manage to carry idiosyncratic information of players into the estimation, while the belief learning models cannot.

to conclude that our model also organizes the experimental data well at an individual level.
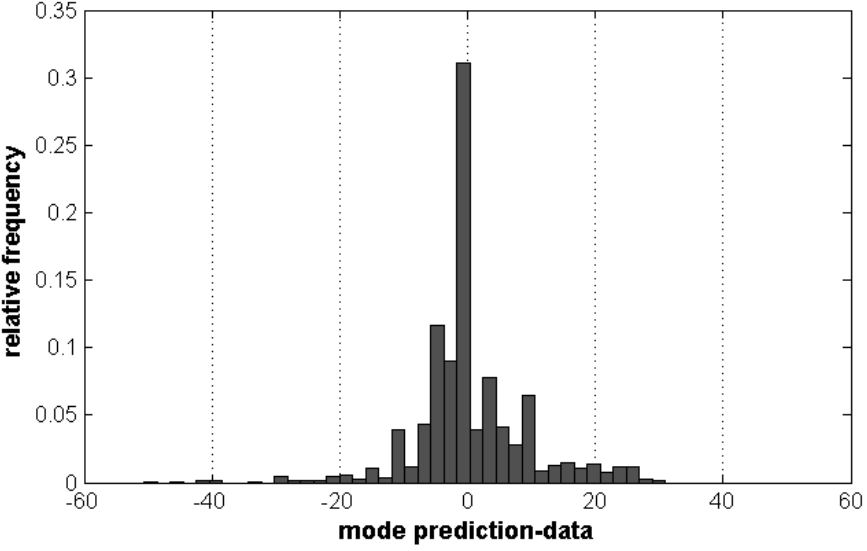

Figure 4.4.3: Histogram for differences between predicted mode prices and posted prices

4.5   Conclusion

We have developed an extended payoff assessment learning model to explain pricing and learning behaviour in experimental Bertrand markets with limited information. While updating payoff assessments, the sellers not only change the assessments of the selected price, but also adjust that associated with other prices. A key difference of our model to the existing learning models is that we explicitly incorporate players' ex-post inferences for the forgone payoffs to unchosen strategies. For some prices, the forgone payoffs can be directly calculated from the feedback received. For the remaining prices, however, direct inference is not feasible. This is because the forgone payoffs for these prices are subjective and determined by sellers' beliefs about the opponents' actions. We showed that a simple parametric approach can help to estimate the subjective forgone payoffs

straight from the data.

An interesting following up of our model is to ask whether the learning process will approach to some stationary distribution of prices over time. For the original payoff assessment model, Sarin & Vahid (1999) proved that if the players only choose the strategies with the highest payoff assessments, and update only the assessments for the chosen strategies, the learning process will converge to the maxmin choices. For the Bertrand duopoly game we study, the maxmin prices are the marginal cost price and thus constitute a Nash equilibrium. However, both of the assumptions that lead to the maxmin result are relaxed in our extended model, which makes the Nash equilibrium unlikely to be reached.

Impulse Balance Equilibrium (Selten & Chmura 2008), which is based on learning directional theory, provides a clue to our context. According to the impulse balance theory, when a player receives feedback, she is assumed to have an impulse to change in the direction of higher payoffs. At an Impulse Balance Equilibrium (IBE), the expected impulses for opposite directions of change are equal. In our environment, we might have the case that a seller's expected impulse to adjust prices upward after a successful sale is balanced with the expected impulse to reduce prices after an unsuccessful sale. Consequently, one can see how an IBE could come about. Simulation or experimental investigation of the model is required in order to test whether the learning process will converge in the long run. And, if the process does converge, whether the distribution of prices are consistent with IBE or other stationary equilibrium concepts.

Bibliography

Abbink, K., Sadrieh, A. & Zamir, S. (2004), 'Fairness, public good, and emotional aspects of punishment behavior', *Theory and decision* **57**(1), 25–57. 44

Abrams, E., Sefton, M. & Yavas, A. (2001), 'An experimental comparison of two search models', *Advances in Experimental Markets* pp. 71–85. 31

Akiva, M. E. B. & Lerman, S. R. (1985), *Discrete choice analysis: theory and application to predict travel demand*, Vol. 9, MIT press. 58

Baye, M. R. & Morgan, J. (2001), 'Information gatekeepers on the internet and the competitiveness of homogeneous product markets', *American Economic Review* **91**, 454–474. 1

Baye, M. R. & Morgan, J. (2004), 'Price dispersion in the lab and on the internet: Theory and evidence', *The RAND Journal of Economics* **35**(3), 449–466. 2, 3, 4, 9, 10, 14, 31, 32, 35

Baye, M. R., Morgan, J. & Scholten, P. (2004), 'Price dispersion in the small and in the large: Evidence from an internet price comparison site', *The Journal of Industrial Economics* **52**(4), 463–496. 31

Bayer, R. C. & Ke, C. (2011), 'Are rockets and feathers caused by search or informational frictions?', *Working Papers* . 2, 5, 31, 45, 49

Bayer, R.-C., Renner, E. & Sausgruber, R. (2013), 'Confusion and learning in the voluntary contributions game', *Experimental Economics* pp. 1–19. 48

Brockwell, P. J. & Davis, R. A. (2009), *Time series: theory and methods*, Springer. 48

Bruttel, L. V. (2009), 'Group dynamics in experimental studies¡ªthe bertrand paradox revisited', *Journal of Economic Behavior & Organization* **69**(1), 51–63. 20, 47

Burdett, K. & Judd, K. L. (1983), 'Equilibrium price dispersion', *Econometrica: Journal of the Econometric Society* pp. 955–969. 31

Bush, R. R. & Mosteller, F. (1955), 'Stochastic models for learning', *Wiley Publications on Statistics* . 44

Camerer, C. F. & Ho, T. H. (1999), 'Experience-weighted attraction learning in normal form games', *Econometrica* **67**(4), 827–874. 48, 59

Camerer, C., Ho, T. & Chong, J. (2004), 'A cognitive hierarchy model of games', *The Quarterly Journal of Economics* **119**(3), 861–898. 59

Capra, C. M., Goeree, J. K., Gomez, R. & Holt, C. A. (2002), 'Learning and noisy equilibrium behavior in an experimental study of imperfect price competition', *International Economic Review* **43**(3). 32

Cason, T. N. & Friedman, D. (1997), 'Price formation in single call markets', *Econometrica: Journal of the Econometric Society* pp. 311–345. 47

Chen, H. C., Friedman, J. W. & Thisse, J. F. (1997), 'Boundedly rational nash equilibrium: a probabilistic choice approach', *Games and Economic Behavior* **18**(1), 32–54. 9, 21

Chen, Y. & Khoroshilov, Y. (2003), 'Learning under limited information', *Games and Economic Behavior* **44**(1), 1–25. 48, 59

Chen, Y. & Tang, F. (1998), 'Learning and incentive-compatible mechanisms for public goods provision: An experimental study', *Journal of Political Economy* **106**(3), 633–662. 59

Cheung, Y. & Friedman, D. (1997), 'Individual learning in normal form games: Some laboratory results', *Games and Economic Behavior* **19**(1), 46–76. 3, 12

Clay, K., Krishnan, R. & Wolff, E. (2001), 'Prices and price dispersion on the web: evidence from the online book industry', *The Journal of Industrial Economics* **49**(4), 521–539. 31

Dufwenberg, M. & Gneezy, U. (2000), 'Price competition and market concentration: an experimental study', *International Journal of Industrial Organization* **18**(1), 7–22. 2, 31

Dufwenberg, M., Gneezy, U., Goeree, J. & Nagel, R. (2007), 'Price floors and competition', *Economic Theory* **33**(1), 211–224. 32, 35, 58

Erev, I. & Roth, A. (1998), 'Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria', *American economic review* pp. 848–881. 44, 59

Goeree, J. K., Holt, C. A. & Palfrey, T. R. (2002), 'Quantal response equilibrium and overbidding in private-value auctions', *Journal of Economic Theory* **104**(1), 247–272. 11

Goeree, J. K., Holt, C. A. & Palfrey, T. R. (2005), 'Regular quantal response equilibrium', *Experimental Economics* **8**(4), 347–367. 13

Grosskopf, B. (2003), 'Reinforcement and directional learning in the ultimatum game with responder competition', *Experimental Economics* **6**(2), 141–158. 47

Ho, T., Camerer, C. & Chong, J. (2007), 'Self-tuning experience weighted attraction learning in games', *Journal of Economic Theory* **133**(1), 177–198. 59

Holt, C. & Laury, S. (2002), 'Risk aversion and incentive effects', *The American economic review* **92**(5), 1644–1655. 11

Huck, S., Normann, H. & Oechssler, J. (2000), 'Does information about competitors¡¯ actions increase or decrease competition in experimental oligopoly markets?', *International Journal of Industrial Organization* **18**(1), 39–57. 44

Janssen, M. C. & Moraga-González, J. L. (2004), 'Strategic pricing, consumer search and the number of firms', *The Review of Economic Studies* **71**(4), 1089–1118. 43

Lach, S. (2002), 'Existence and persistence of price dispersion: an empirical analysis', *Review of Economics and Statistics* **84**(3), 433–444. 31

López-Acevedo, G. (1997), 'Quantal response equilibria for posted offer-markets', *Estudios Económicos* pp. 95–131. 9

Luce, R. (1959), *Individual choice behavior.*, John Wiley. 9, 32

Maskin, E. & Tirole, J. (1988), 'A theory of dynamic oligopoly, ii: Price competition, kinked demand curves, and edgeworth cycles', *Econometrica: Journal of the Econometric Society* pp. 571–599. 21

McFadden, D. (1973), 'Conditional logit analysis of qualitative choice behavior'. 32

McKelvey, R. D. & Palfrey, T. R. (1995), 'Quantal response equilibria for normal form games', *Games and Economic Behavior* **10**(1), 6–38. 2, 31, 58

Mookherjee, D. & Sopher, B. (1994), 'Learning behavior in an experimental matching pennies game', *Games and Economic Behavior* **7**(1), 62–91. 44

Radner, R. (1980), 'Collusive behavior in noncooperative epsilon-equilibria of oligopolies with long but finite lives', *Journal of Economic Theory* **22**(2), 136–154. 2

Rauh, M. (2001), 'Heterogeneous beliefs, price dispersion, and welfare-improving price controls', *Economic Theory* **18**(3), 577–603. 2

Reinganum, J. F. (1979), 'A simple model of equilibrium price dispersion', *The Journal of Political Economy* pp. 851–858. 1, 31

Rosenthal, R. W. (1980), 'A model in which an increase in the number of sellers leads to a higher price', *Econometrica: Journal of the Econometric Society* pp. 1575–1579. 43

Rosenthal, R. W. (1989), 'A bounded-rationality approach to the study of noncooperative games', *International Journal of Game Theory* **18**(3), 273–292. 9

Salop, S. & Stiglitz, J. (1977), 'Bargains and ripoffs: A model of monopolistically competitive price dispersion', *The Review of Economic Studies* **44**(3), 493–510. 1

Sarin, R. & Vahid, F. (1999), 'Payoff assessments without probabilities: a simple dynamic model of choice', *Games and Economic Behavior* **28**(2), 294–309. iii, 23, 25, 30, 67

Sarin, R. & Vahid, F. (2001), 'Predicting how people play games: a simple dynamic model of choice', *Games and Economic Behavior* **34**(1), 104–122. 44

Sarin, R. & Vahid, F. (2004), 'Strategy similarity and coordination*', *The Economic Journal* **114**(497), 506–527. 48

Satterthwaite, M. A. (1979), 'Consumer information, equilibrium industry price, and the number of sellers', *The Bell Journal of Economics* pp. 483–502. 43

Selten, R., Abbink, K. & Cox, R. (2005), 'Learning direction theory and the winner¡¯s curse', *Experimental Economics* **8**(1), 5–20. 47

Selten, R. & Buchta, J. (1999), 'Experimental sealed bid first price auctions with directly observed bid functions', *Games and human behavior: Essays in honor of Amnon Rapoport* pp. 101–116. 47

Selten, R. & Chmura, T. (2008), 'Stationary concepts for experimental 2x2-games', *The American Economic Review* **98**(3), 938–966. 67

Selten, R. & Stoecker, R. (1986), 'End behavior in sequences of finite prisoner's dilemma supergames a learning theory approach', *Journal of Economic Behavior & Organization* **7**(1), 47–70. 47

Stigler, G. J. (1961), 'The economics of information', *The Journal of Political Economy* **69**(3), 213–225. 1

Stiglitz, J. E. (1987), 'Competition and the number of firms in a market: Are duopolies more competitive than atomistic markets?', *The Journal of Political Economy* **95**(5), 1041–1061. 43

Varian, H. R. (1980), 'A model of sales', *The American Economic Review* **70**(4), 651–659. 1, 31, 43

Wilcox, N. (2006), 'Theories of learning in games and heterogeneity bias', *Econometrica* **74**(5), 1271–1292. 65