Identification of Firms' Beliefs
in Structural Models of Market Competition

Victor Aguirregabiria*
University of Toronto and CEPR

December 26th, 2019

Abstract

This paper studies the joint identification of firms' beliefs and structural parameters in a general class of empirical games of market competition. In this framework, firms have incomplete information and their beliefs about the behavior of other firms are unrestricted – nonparametric – functions of a firm's information. Beliefs may be out of equilibrium. This framework includes — as particular cases — models of competition in prices or quantities, auction models, and static and dynamic discrete choice games. I present identification results under different scenarios on the data available to the researcher: only firms' choice data; data on consumers' demand; and data on firms' costs.

Keywords: Incomplete information; Non-equilibrium beliefs; Oligopoly competition; Structural models; Identification.

JEL codes: C57, D81, D83, D84, L13.

Victor Aguirregabiria. Address: 150 St. George Street. Toronto, ON, M5S 3G7, Canada. Phone: (416) 978-4358. E-mail: victor.aguirregabiria@utoronto.ca

---

1 Introduction

A firm's behavior depends on its beliefs about the actions of other firms in the same market. For instance, the optimal price of a firm depends on its beliefs about the prices of its competitors; or, in a procurement auction, a firm's bid depends on its expectation about other firms' bids. Managers form their beliefs under uncertainty about demand, costs, and competitors' decisions. For similar reasons as they are heterogeneous in their efficiency to produce goods and services, firms are different in their ability to collect and process information. As a result, we expect firms to be heterogeneous in their expectations or beliefs, and this heterogeneity has implications on their performance and on their survival in the market. For the same reasons as firms with different productivity can survive in the same industry, they can coexist with different levels of accuracy or bias in their beliefs.

The importance of firms' heterogeneity in their ability to form expectations — and the possibility of biased or non-equilibrium beliefs — has been long recognized in economics, at least since the work of Simon (1958, 1959). However, in most fields in economics, the status quo has been to assume rational expectations. In empirical industrial organization (IO), some of the most commonly used structural models of oligopoly competition assume complete information, perfect certainty, and Nash equilibrium. For instance, this is the case in models of price competition with differentiated product (Berry, Levinsohn, and Pakes, 1995; Berry and Haile, 2014), and in empirical games of market entry (Bresnahan and Reiss, 1991; Ciliberto and Tamer, 2009). Though there is a substantial literature on structural models of incomplete information in empirical IO, it is mostly concentrated in auctions (Guerre, Perrigne, and Vuong, 2000; Athey and Haile, 2002), and in discrete choice games, both static (Seim, 2006; Sweeting, 2009), and dynamic (Aguirregabiria and Mira, 2007; Igami, 2017). Most empirical applications of games of incomplete information assume that firms have homogeneous beliefs that correspond to a Bayesian Nash equilibrium.

Recent applications in empirical IO relax equilibrium assumptions and obtain evidence of heterogeneity and biases in firms' beliefs. As one would expect, biased beliefs are more likely in new markets and after regulatory changes: for instance, after deregulation of the US telecommunication industry (Goldfarb and Xiao, 2011), the UK electricity market (Doraszelski, Lewis, and Pakes, 2018), or the Washington State liquor market (Huang, Ellickson, and Lovett, 2018), and in the early years of the fast-food restaurant industry in UK (Aguirregabiria and Magesan, 2019). These empirical studies use different approaches to identify firms' beliefs.

This paper presents a systematic analysis of the joint identification of firms' beliefs and struc-

tural parameters in a general class of empirical games of market competition. The main emphasis is that — under weak restrictions — firms' observed behavior reveals information about their beliefs.

I consider a framework where firms have incomplete information. A firm's beliefs about the behavior of other firms are unrestricted – nonparametric – functions of a firm's information. Beliefs may be out of equilibrium. This framework includes — as particular cases — models of competition in prices or quantities, auction models, and static and dynamic discrete choice games. I present identification results under different scenarios on the data available to the researcher: from data only on firms' choices and state variables, to situations where the researcher can identify the revenue function, or even the cost function. The identification results vary substantially when the model of competition is static or dynamic, so I present these two types of results.

Though I focus on the possibility of non-equilibrium beliefs, the results in this paper also apply to the identification of collusive behavior under asymmetric information. As such, this paper relates to the traditional literature on identification of the form or nature of competition, pioneered by Bresnahan (1982, 1987) and Nevo (1998, 2001).

Section 2 presents a brief description of recent applications that provide evidence on firms' biased beliefs and their important implications for market outcomes. In section 3, I present a general framework that includes as particular cases almost every model of competition that we find in empirical applications in IO. A key feature of this framework is that it allows for a very flexible (nonparametric) specification of firms' probabilistic beliefs as functions of firms' information. Section 4 presents the results on the identification of beliefs. I conclude in section 5 with a discussion of different issues such as the direct measurement of firms' beliefs from survey data, and the implementation of counterfactual experiments in models with biased beliefs.

2   Empirical evidence on firms' biased beliefs

In this section, I briefly describe the empirical evidence in recent studies that find evidence on firms' biased or non-equilibrium beliefs. All these studies estimate structural models of oligopoly competition. Here, I first describe the reduced form evidence in each of these papers that motivates relaxing the assumption of equilibrium.

Goldfarb and Xiao (2011) study entry decisions into local US telecommunication markets following the deregulatory Telecommunications Act of 1996, which allowed free competition. The authors present reduced-form evidence showing that, holding other market characteristics constant, more experienced and better educated managers have a lower propensity to enter (and a lower propensity

to exit after entry) into very competitive markets. This suggests that better-educated managers are better at predicting competitors' behavior.

Aguirregabiria and Magesan (2019) study competition in store location between McDonalds (MD) and Burger King (BK) during the early years of the fast-food restaurant industry in the UK. Estimates of reduced form models for the decision of opening a new store show that the number of own stores has a strong negative effect on the probability that BK opens a new store but the effect of the competitor's number of stores is negligible. In contrast, for MD, the decision to open a new store is sensitive to BK's existing number of stores in the market. Standard equilibrium models of market entry — where firms compete and sell substitute products — cannot rationalize this behavior.

Doraszelski, Lewis, and Pakes (2018) investigate firms' learning about competitors' bidding behavior just after the deregulation of the UK electricity market. In the first months (year) after deregulation, firms' bidding behavior was very heterogeneous and firms made frequent and sizable adjustments in their bids. During the next year, there was a dramatic reduction in the range of bids; and after three years, firms' bids became very stable. During these three phases, demand and costs were very stable. The authors argue that the changes in firms' bidding strategies can be attributed to strategic uncertainty and learning rather than changes in the environment.

Huang, Ellickson, and Lovett (2018) study firms' price setting behavior in the Washington State liquor market following the privatization of the market in 2012. After liberalization, grocery chains newly entered the market. How did these new entrants learn about demand and learn to price optimally over time? The authors document that there are large and heterogeneous price movements in the first two years after the privatization. The authors present evidence that is consistent with firms' learning about the idiosyncratic and common components of the demand shocks and about the time persistence of these shocks. The estimation of a structural model and counterfactual experiments reveal a 13% loss in the profits of inexperienced sellers due to the information frictions.

Hortaçsu and Puller (2008) analyze the bidding behavior of firms in the Texas electricity spot market, where suppliers trade with each other. Their dataset contains detailed information not only on firms' bids but also on their marginal costs. Using these data on marginal costs, the authors construct the equilibrium bids of the game and compare them to the actual observed bids. They find statistically and economically very significant deviations between equilibrium and actual bids. While large firms best-respond to other firms' behavior, small firms submit bid-functions that

3

are excessively steep. Small firms supply little power even when it is ex-post profitable to do so. Based on their interviews with traders in this market, Hortaçsu and Puller argue that this finding is best explained by the relatively low strategic ability in the bidding departments of small firms. Very interestingly, this suboptimal behavior by small firms leads to significant efficiency losses in the market. In fact, the inefficiency generated by the mistakes of smaller firms is larger than the inefficiency generated by the market power of large firms.

3   Model

3.1   Static game

3.1.1   Basic framework

Consider N firms competing in a market. Firms are indexed by i. The profit function of firm i is $\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x})$ where $a_i \in \mathcal{A}$ is the action of firm i, $\boldsymbol{a}_{-i} = (a_j : j \neq i) \in \mathcal{A}^{N-1}$ is the vector with the actions of the other firms. This is a general model of market competition where a firm's decision variable – a_i – can be either discrete or continuous, and it can represent – among other possibilities – a firm's level of output, its price, the binary indicator of entry in a market, the firm's number of stores, or its investment in R&D. The vector $\mathbf{x} \in \mathcal{X}$ represents variables that are common knowledge to all the players. The term $\varepsilon_i \in \mathcal{E}$ is private information of firm i. For instance, this private information can be a component of the firm's cost or a private signal about the state of the demand. We denote ε_i as the 'type' of firm i, which is privately known. Firms' types $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_N)$ are drawn from a distribution with cumulative distribution function $F(\varepsilon_1, \varepsilon_2, ..., \varepsilon_N \mid \mathbf{x})$.

Firms simultaneously decide their actions to maximize their respective expected profits. Under the standard solution concept of Bayesian Nash Equilibrium (BNE), the primitives of the model – that is π's, $\mathcal{A}$, $\mathcal{X}$, $\mathcal{E}$, and F – are assumed common knowledge – that is, every firm knows that every firm knows ... these primitives. The model that I consider here does not impose this restriction. This model only assumes that each firm knows its own profit function, the vector of variables $\mathbf{x}$, and its private information ε_i. For instance, some firms may not know the distribution function F or the profit functions $\pi_i(.)$ of other firms. Furthermore, firms may not know that $\mathbf{x}$ is common knowledge.

A firm does not know the private information of its competitors and therefore it does not know their actions. Firms form probabilistic beliefs about the actions of competitors. Let $B_i(\boldsymbol{a}_{-i} \mid \varepsilon_i, \mathbf{x})$ be a probability density function that represents the belief of firm i. This is a probability function in the space of the actions of firms other than i and conditional on firm $i's$ information. Given its

4

beliefs, a firm's expected profit is:

$$\pi_i^e(a_i, \varepsilon_i, \mathbf{x}, B_i) = \int_{\boldsymbol{a}_{-i}} \pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i}|\varepsilon_i, \mathbf{x}) \, d\boldsymbol{a}_{-i}, \tag{1}$$

The integral is over the Lebesgue measure on $\mathcal{A}^{N-1}$, which can be either continuous or discrete.

A firm chooses its strategy function $\sigma_i(\varepsilon_i, \mathbf{x}, B_i)$ to maximize expected profits:

$$\sigma_i(\varepsilon_i, \mathbf{x}, B_i) = \arg \max_{a_i \in \mathcal{A}} \pi_i^e(a_i, \varepsilon_i, \mathbf{x}, B_i). \tag{2}$$

It is convenient to represent a firm's strategy as a cumulative distribution function. Let $P_i(a_i|\mathbf{x})$ be cumulative distribution of the choice variable a_i conditional on $\mathbf{x}$. We denote $P_i(a_i|\mathbf{x})$ as the *cumulative choice probability function*. According to the model, this distribution comes from firm i's best response and satisfies the following equation. For any value $a^0 \in \mathcal{A}$,

$$P_i\left(a^0|\mathbf{x}\right) \equiv \Pr\left(\sigma_i(\varepsilon_i, \mathbf{x}, B_i) \leq a^0 \mid \mathbf{x}; B_i\right) = \int 1\left\{\sigma_i(\varepsilon_i, \mathbf{x}, B_i) \leq a^0\right\} \, dF_i(\varepsilon_i|\mathbf{x}) \tag{3}$$

In this framework, the payoff functions $\{\pi_i\}$ and the distribution of private signals, $F(.|\mathbf{x})$, are primitives of the model. The predictions of the model are the choice probabilities. The belief functions $\{B_i\}$ are endogenous outcomes of the model. However, the model is incomplete in the sense that it does not specifies how these beliefs are determined. Instead, it specifies a general framework that includes as particular cases many different models for the determination of beliefs.

3.1.2 Characterization of best response functions

I focus in models where a firm's action a_i is a single variable that can be either continuous or discrete. If the decision is discrete, then it is ordered, e.g., number of products, stores, etc. Note that any binary choice (e.g., a market entry decision) can be considered a particular case of ordered discrete choice variable. The framework also imposes some restrictions on the marginal profit function such that a firm's best response function is strictly monotonic in ε_i. The following paragraphs describe these assumptions.

Let $\Delta\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x})$ be the *marginal profit function*. Here the concept of *marginal profit* is broad and depends on the decision variables a_i. If the decision variable is continuous – such as output or price – the marginal profit is at the *intensive margin* and it corresponds to the mathematical concept of the derivative of a continuous and differentiable function: $\Delta\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) = \partial\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x})/\partial a_i$. If the decision variable is discrete – such as entry decision, number of stores, or number of products – the marginal profit is at the *extensive margin* and it corresponds to the difference function: $\Delta\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) = \pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) - \pi_i(a_i - 1, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x})$.

The first condition of optimality for the best response of firm i is: (i) for continuous choice variable, $\Delta \pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) = 0$; and (ii) for discrete choice, $\Delta \pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) \geq 0$ and $\Delta \pi_i^e(a_i + 1, \varepsilon_i, \mathbf{x}; B_i) \leq 0$. Taking into account the definition of the expected profit, we can represent these conditions as follows.

(i) Continuous choice variable ($a_i \in \mathbb{R}$):

$$\int_{\boldsymbol{a}_{-i}} \Delta \pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i} | \varepsilon_i, \mathbf{x}) \, d\boldsymbol{a}_{-i} = 0 \qquad (4)$$

(ii) Ordered discrete choice variable ($a_i \in \{0, 1, ..., J\}$):

$$\begin{cases} a_i = 0 & \Leftrightarrow \quad \sum_{\boldsymbol{a}_{-i}} \Delta \pi_i(1, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i} | \varepsilon_i, \mathbf{x}) < 0 \\ & \qquad\qquad \sum_{\boldsymbol{a}_{-i}} \Delta \pi_i(j, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i} | \varepsilon_i, \mathbf{x}) \geq 0 \\ a_i = j \text{ for } 0 < j < J & \Leftrightarrow \quad \text{and} \ \sum_{\boldsymbol{a}_{-i}} \Delta \pi_i(j+1, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i} | \varepsilon_i, \mathbf{x}) < 0 \\ a_i = J & \Leftrightarrow \quad \sum_{\boldsymbol{a}_{-i}} \Delta \pi_i(J, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i} | \varepsilon_i, \mathbf{x}) \geq 0 \end{cases} \qquad (5)$$

The following assumptions provide sufficient conditions for the strict monotonicity of a firm's best response function with respect to its type ε_i.

ASSUMPTION 1. For any value of $(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x})$, the marginal profit function $\Delta \pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x})$ is: (A) strictly decreasing in the action variable a_i; and (B) strictly monotonic – either increasing or decreasing – in ε_i. ∎

Assumption 1 holds in most models of market competition. Suppose that we represent a firm's profit as revenue (R_i) minus cost (C_i) such that $\pi_i = R_i - C_i$. For models of competition in price or quantity, a downward sloping demand curve and non-decreasing marginal costs are sufficient conditions for the concavity of the profit function in a_i. A sufficient condition for assumption 1(B) is that the type variable ε_i enters only in the cost function and the marginal cost is monotonic in ε_i – e.g., a firm's type captures cost efficiency that is private information of the firm. Assumption 1 also holds in models of procurement auctions and in market entry models where the firm's type is a component of its cost.

A particular case of Assumption 1(B) is when the profit function is additive in ε_i. That is,

$$\Delta \pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) = \Delta \pi_i^*(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) - \varepsilon_i. \qquad (6)$$

This condition is stronger than Assumption 1(B) but it is useful to provide a simpler version of some of the results in this paper.

Assumption 2 establishes the restriction that firms' types are independently distributed – *independent private values (IPV)*. This restriction is stronger than what we need to characterize best response functions and to obtain the identification results in this paper. However, it is convenient because it facilitates the derivations and proofs in this paper. I comment below on weaker assumptions on the joint distribution of the private information variables.

ASSUMPTION 2. (A) The private information variables $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_N)$ are independently distributed conditional on $\mathbf{x}$*:* $F(\varepsilon_1, \varepsilon_2, ..., \varepsilon_N | \mathbf{x}) = \prod_{i=1}^{N} F_i(\varepsilon_i | \mathbf{x})$*. (B) Every firm i knows this independence, and therefore knows that* $\boldsymbol{a}_{-i}$ *and* ε_i *are independent conditional on* $\mathbf{x}$*.* ■

Assumption 2 implies that ε_i is not an argument of the beliefs function $B_i(\boldsymbol{a}_{-i} | \varepsilon_i, \mathbf{x})$. Every firm is aware that – conditional on $\mathbf{x}$ – the actions of the other firms are independent of its own type. Firms have beliefs that are consistent with this independence. For the rest of the paper, I represent a belief function as $B_i(\boldsymbol{a}_{-i} | \varepsilon_i, \mathbf{x})$.

ASSUMPTION 3. The private information ε_i is a continuous random variable with support the real line and with a cumulative distribution function conditional on $\mathbf{x}$ *–* $F_i(\varepsilon_i | \mathbf{x})$ *– that is strictly increasing over the whole real line.* ■

Assumption 1(A) implies that the expected marginal profit – $\Delta \pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i)$ – is strictly decreasing in a_i. By the *Implicit Function Theorem*, assumption 1(A) implies that equation $\Delta \pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) = 0$ is invertible with respect $a_i \in \mathbb{R}$. That is, for any value of $(\varepsilon_i, \mathbf{x}, B_i)$, there is a unique value of $a_i \in \mathbb{R}$ – that I denote as $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ – such that $\Delta \pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) = 0$ if and only if $a_i = a_i^*(\varepsilon_i, \mathbf{x}, B_i)$.

If the decision variable is continuous, then $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ is the best response function of firm i – that is, $\sigma_i(\varepsilon_i, \mathbf{x}, B_i) = a_i^*(\varepsilon_i, \mathbf{x}, B_i)$. If the decision variable is discrete, then in general $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ does not belong to the discrete choice set $\mathcal{A}$ such that $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ is not the optimal decision rule. However, by construction and by monotonicity of the profit function, the largest integer less than or equal to $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ is such that its expected profit is greater or equal than zero; and the next larger integer has an expected profit smaller or equal than zero. Therefore, the optimal decision rule is the *floor function* of $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ – that is, $\sigma_i(\varepsilon_i, \mathbf{x}, B_i) = \lfloor a_i^*(\varepsilon_i, \mathbf{x}, B_i) \rfloor$.[1] Given

---

[1] Remember that the floor function $\lfloor x \rfloor$ is defined as the largest integer less than or equal to x.

that $\Delta\pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i)$ is strictly decreasing in a_i, we have that $\Delta\pi_i^e(\lfloor a_i^*(\varepsilon_i, \mathbf{x}, B_i)\rfloor], \varepsilon_i, \mathbf{x}; B_i) \geq 0$ and $\Delta\pi_i^e(\lfloor a_i^*(\varepsilon_i, \mathbf{x}, B_i)\rfloor + 1, \varepsilon_i, \mathbf{x}; B_i) \leq 0$ such that $\lfloor a_i^*(\varepsilon_i, \mathbf{x}, B_i)\rfloor$ is the best response function.

As defined above, the *cumulative choice probability function* $P_i\left(a^0 \mid \mathbf{x}; B_i\right)$ is the cumulative distribution for a_i implied by the best response function $\sigma_i(\varepsilon_i, \mathbf{x}, B_i)$ – that is, $P_i(a^0|\mathbf{x}) = \int 1\{\sigma_i(\varepsilon_i, \mathbf{x}; B_i) \leq a^0\}\, dF_i(\varepsilon_i|\mathbf{x})$. Proposition 1 establishes that function $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ is strictly monotonic in ε_i and this implies a simple and convenient expression for the *cumulative choice probability function.*

PROPOSITION 1. Under assumptions 1 to 3, function $a_i^*(\varepsilon_i, \mathbf{x}, B_i)$ is strictly monotonic in ε_i – say, strictly increasing, without loss of generality. This implies that, for any $a^0 \in \mathcal{A}$:

$$P_i(a^0|\mathbf{x}) = F_i\left(e_i^*(a^0, \mathbf{x}, B_i) \mid \mathbf{x}\right) \tag{7}$$

where $e_i^*(a^0, \mathbf{x}, B_i)$ is the inverse function that results from applying the inverse function theorem to equation $\Delta\pi_i^e(a^0, \varepsilon_i, \mathbf{x}; B_i) = 0$ with respect to ε_i. ∎

The following examples illustrate Proposition 1 using four different types of models of competition: Cournot, auctions, market entry, and number of stores, respectively.

EXAMPLE 1 (Cournot competition). Consider a game of quantity choice where the output of firm i is $a_i \in \mathbb{R}_+$. Suppose that the marginal profit function is additive in ε_i – e.g., ε_i is an additive component of the marginal cost: that is, $\Delta\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) = \Delta\pi_i^*(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) - \varepsilon_i$. Then, the expected marginal profit function is:

$$\Delta\pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) = \int_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(a_i, \boldsymbol{a}_{-i}, \mathbf{x})\, B_i(\boldsymbol{a}_{-i}|\mathbf{x})\, d\boldsymbol{a}_{-i} - \varepsilon_i \tag{8}$$

Therefore, by definition of function e_i^*, we have that $e_i^*(a^0, \mathbf{x}, B_i) = \int_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(a^0, \boldsymbol{a}_{-i}, \mathbf{x})\, B_i(\boldsymbol{a}_{-i}|\mathbf{x})\, d\boldsymbol{a}_{-i}$, and the cumulative choice probability function has the following form:

$$P_i(a^0|\mathbf{x}) = F_i\left(\int_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(a^0, \boldsymbol{a}_{-i}, \mathbf{x})\, B_i(\boldsymbol{a}_{-i}|\mathbf{x})\, d\boldsymbol{a}_{-i} \mid \mathbf{x}\right). \quad ∎ \tag{9}$$

EXAMPLE 2 (Procurement auction). Consider a procurement auction where $a_i \in \mathbb{R}$ represents firm i's bid. The profit function is $\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) = (a_i - c_i(\varepsilon_i, \mathbf{x}))\, 1\{a_j > a_i\ \forall j \neq i\}$, where

$c_i(\varepsilon_i, \mathbf{x})$ is the cost function. The expected profit function is:

$$\pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) \;=\; (a_i - c_i(\varepsilon_i, \mathbf{x})) \int_{\boldsymbol{a}_{-i}} 1\{a_j > a_i \ \forall j \neq i\} \ B_i(\boldsymbol{a}_{-i}|\mathbf{x}) \ d\boldsymbol{a}_{-i} \tag{10}$$

$$=\; (a_i - c_i(\varepsilon_i, \mathbf{x})) \ W(a_i, \mathbf{x}, B_i)$$

where $W(a_i, \mathbf{x}, B_i) \equiv \int_{\boldsymbol{a}_{-i}} 1\{a_j > a_i \ \forall j \neq i\} \ B_i(\boldsymbol{a}_{-i}|\mathbf{x}) \ d\boldsymbol{a}_{-i}$ is firm i's subjective probability of wining the auction given its beliefs. Therefore, expected marginal profit function is:

$$\Delta\pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) \;=\; W(a_i, \mathbf{x}, B_i) + (a_i - c_i(\varepsilon_i, \mathbf{x})) \ \Delta W(a_i, \mathbf{x}, B_i), \tag{11}$$

where $\Delta W(a_i, \mathbf{x}, B_i)$ represents the derivative of the subjective probability of winning with respect to the own bid a_i. Suppose that the cost function $c_i(\varepsilon_i, \mathbf{x})$ is additive in ε_i: that is, $c_i(\varepsilon_i, \mathbf{x}) = c_i^*(\mathbf{x}) + \varepsilon_i$. Then, we have that

$$e_i^*(a^0, \mathbf{x}, B_i) = a^0 - c_i^*(\mathbf{x}) + \frac{W(a^0, \mathbf{x}, B_i)}{\Delta W(a^0, \mathbf{x}, B_i)} \tag{12}$$

And the cumulative choice probability function is:

$$P_i(a^0|\mathbf{x}) = F_i\left(a^0 - c_i^*(\mathbf{x}) + \frac{W(a^0, \mathbf{x}, B_i)}{\Delta W(a^0, \mathbf{x}, B_i)} \ | \ \mathbf{x} \right). \qquad \blacksquare \tag{13}$$

EXAMPLE 3 (Market entry game). Consider a game of market entry with $a_i \in \{0, 1\}$. Suppose that the profit function is additive in ε_i – e.g., ε_i is a component of the entry cost or fixed cost: that is, $\Delta\pi_i(1, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \equiv \Delta\pi_i^*(1, \boldsymbol{a}_{-i}, \mathbf{x}) - \varepsilon_i$. Then, the expected marginal profit function is:

$$\Delta\pi_i^e(1, \varepsilon_i, \mathbf{x}; B_i) \;=\; \sum_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(1, \boldsymbol{a}_{-i}, \mathbf{x}) \ B_i(\boldsymbol{a}_{-i}|\mathbf{x}) - \varepsilon_i \tag{14}$$

Therefore, by definition of function e_i^*, we have that $e_i^*(1, \mathbf{x}, B_i) = \sum_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(1, \boldsymbol{a}_{-i}, \mathbf{x}) \ B_i(\boldsymbol{a}_{-i}|\mathbf{x})$, and the choice probability function is:

$$P_i(0|\mathbf{x}) = 1 - F_i\left(\sum_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(1, \boldsymbol{a}_{-i}, \mathbf{x}) \ B_i(\boldsymbol{a}_{-i}|\mathbf{x}) \ | \ \mathbf{x} \right). \qquad \blacksquare \tag{15}$$

EXAMPLE 4 (Competition in number of stores or products). Consider a game of market entry $a_i \in \{0, 1, ..., J\}$ represents the number of products that the firm has in the market. Suppose that the marginal profit function is additive in ε_i – e.g., ε_i is a component of the marginal cost of introducing a new product: that is, $\Delta\pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) \equiv \Delta\pi_i^*(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) - \varepsilon_i$. Therefore, the expected marginal profit function is $\Delta\pi_i^e(a_i, \varepsilon_i, \mathbf{x}; B_i) = \sum_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) \ B_i(\boldsymbol{a}_{-i}|\mathbf{x}) - \varepsilon_i$, and we have that $e_i^*(a^0, \mathbf{x}, B_i) = \sum_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(a^0, \boldsymbol{a}_{-i}, \mathbf{x}) \ B_i(\boldsymbol{a}_{-i}|\mathbf{x})$. For $0 < a^0 < J$, the choice probability function is:

$$P_i(a^0|\mathbf{x}) = F_i\left(\sum_{\boldsymbol{a}_{-i}} \Delta\pi_i^*(a^0, \boldsymbol{a}_{-i}, \mathbf{x}) \ B_i(\boldsymbol{a}_{-i}|\mathbf{x}) \ | \ \mathbf{x} \right). \qquad \blacksquare \tag{16}$$

3.1.3 Common equilibrium restrictions in empirical applications

The framework presented above contains as particular cases most of the games of competition with incomplete information that have been considered in empirical applications in IO. The main difference is that in most empirical applications firms' beliefs are assumed to satisfy some equilibrium restrictions. Different equilibrium concepts have been used in the literature. I present here the equilibrium concepts that have received more attention in empirical applications in IO.

All these equilibrium concepts assume that firms choose their best response strategies given their beliefs: that is, they impose the best response conditions described above. In addition, they restrict beliefs to satisfy some additional equilibrium restrictions. I describe below these additional restrictions.

(a) **Bayesian Nash Equilibrium (BNE) with independent private values.** This is the most commonly used solution concept in games of incomplete information in IO. It has received particular attention in auction games (e.g., Guerre, Perrigne, and Vuong, 2000; Athey and Haile, 2002), and in discrete choice games (e.g., Seim, 2006; Sweeting, 2009). It has been used also in empirical applications of Cournot competition models with incomplete information (Armantier and Richard, 2003; Aryal and Zicenko, 2019).

A BNE can be described as N cumulative choice probability functions – $\{P_i(a_i|\mathbf{x}) : i = 1, 2, ..., N\}$ – satisfying the following conditions: (i) [best responses] $P_i(a_i|\mathbf{x})$ satisfies the best response condition given beliefs B_i; and (ii) [rational beliefs] the beliefs function B_i is equal to the actual probability distribution of the choices of the other firms conditional on $\mathbf{x}$:

$$B_i(\boldsymbol{a}_{-i}|\mathbf{x}) = \prod_{j \neq i} \Delta P_j(a_j|\mathbf{x}) \tag{17}$$

where $\Delta P_j(a_j|\mathbf{x})$ represents the density probability function associated to the cumulative distribution function $P_j(a_j|\mathbf{x})$.

(b) **Cognitive Hierarchy (CH) and Level-K models.** These models propose equilibrium concepts where firms have biased beliefs, that is, $B_i(\boldsymbol{a}_{-i} \mid \mathbf{x}) \neq \prod_{j \neq i} \Delta P_j(a_j|\mathbf{x})$. They are based on the following ideas. Firms are heterogeneous in their beliefs and there is a finite number of belief types: that is, $B_i(\boldsymbol{a}_{-i}|\mathbf{x})$ belongs to a finite family of K belief functions, $\{B^{(k)}(\boldsymbol{a}_{-i}|\mathbf{x}) : k = 1, 2, ..., K\}$. Each of member of this family is a 'belief type'. Belief types correspond to different levels of strategic sophistication and are determined by a hierarchical structure.

A firm type-0 has arbitrary believes $B^{(0)}$. In the Level-k model, a type-k firm believes that all the other firms are type-(k-1). Therefore, a type-k firm has beliefs:

$$B^{(k)}(\boldsymbol{a}_{-i}|\mathbf{x}) = \prod_{j\neq i}\Delta F_j\left(e_j^*(a_j, \mathbf{x}, B^{(k-1)}) \mid \mathbf{x}\right) \tag{18}$$

This recursive equation defines the belief functions for every type k between 1 and K. Note that the only unrestricted function is the beliefs function for type-0: the rest of the belief functions are known functions of $B^{(0)}$ and the primitives of the model.

The CH model is more flexible than the Level-k model. In the CH model, a type-k firm believes that the other firms come from a probability distribution over types 0 to (k-1).

These models allows for some flexibility in beliefs. However, they still impose important restrictions. More specifically, they do not include BNE or rational beliefs as a particular case; and there is a small number of belief types – K is smaller than N, and typically 2 or 3 in actual applications.

(c) **Rationalizability** (Bernheim, 1984; Pearce, 1984). The concept of rationalizability imposes two simple restrictions on firms' beliefs and behavior. First, every firm maximizes its own expected profit given beliefs. And second, this rationality is common knowledge, i.e., every firms knows that every firm knows ... that all the firms are rational. The set of outcomes of the game that satisfy these conditions – the set of rationalizable outcomes – includes all the Bayesian Nash equilibria of the game. In a game with multiple equilibria, the solution concept of Rationalizability allows for biased beliefs. Each firm has beliefs that are consistent with a BNE, but these beliefs may not correspond to the same BNE.

In general, the set of rationalizable beliefs is substantially smaller than the set of all the possible beliefs. Therefore, condition (ii) of common knowledge rationality imposes non trivial restrictions with respect to the model that I consider in this paper. In section 4, I show that these additional restrictions are testable.

(d) **Correlated Bayesian Nash Equilibrium.** In recent work, Bergemann and Morris (2013, 2016) have introduced the Bayesian Correlated Equilibrium (BCE) as a solution concept which is more robust, in the sense that it delivers all predictions compatible with Bayesian Nash equilibria for any information structure within a wide class. Magnolfi and Roncoroni (2017) study inference based on the BCE solution concept. Their goal is to identify only the payoff parameters, and their work illustrates a trade-off between robustness to assumptions about information structures and the ability to achieve point identification.

In section 4, I present conditions for the identification of firms' beliefs functions given the framework in section 3.1 and under very weak restrictions on the primitives. Given the identification of beliefs, it is possible to test the additional equilibrium restrictions imposed by the models (a) to (d) described above.

3.2 Dynamic games

In this section, I extend the previous framework to a dynamic game. Time is discrete and indexed by t. Now, $\pi_{it}(a_{it}, \boldsymbol{a}_{-it}, \varepsilon_{it}, \mathbf{x}_t)$ represents the profit function of firm i at period t. The arguments of this function have the same interpretation as before. Firms choose their actions at every period t to maximize their expected and discounted profits $\mathbb{E}_t(\sum_{s=0}^{T-t} \beta^s \ \pi_{it+s})$, where β is the discount factor and T is the time horizon that can be finite or infinite.

I introduce two additional assumptions. First, the common knowledge state variables follow a controlled Markov process with transition probability density function $f_{xt}(\mathbf{x}_{t+1} \mid a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t)$. Second, the private information variables $\varepsilon'_{it}s$ are independently distributed over time.

The restriction that the private information variables are not serially correlated is far of being innocuous. It rules out the possibility of firms using the history of other firms' decisions to learn about these firms' type. This type of learning is the focus of the Experience-Based equilibrium concept in Fershtman and Pakes (2012). In section 4.5, I explain how to extend the framework in this paper to incorporate serially correlated private information.

Every period t, firms select simultaneously their actions to maximize their respective intertemporal values. A firm's value at period t depends on the actions of other firms at period and in the future. A firm does not know the private information of its competitors, now or in the future, and therefore it does not know their actions. Firms form probabilistic beliefs about the actions of competitors, now and in the future.

Let $B_{it+s}^{(t)}(\boldsymbol{a}_{-i,t+s} \mid \mathbf{x}_{t+s})$ be a probability density function, in the space of the actions of firms other than i, that represents the beliefs of firm i at period t about the behavior of other players at period $t + s$. This representation of beliefs is very general and allows for general forms of beliefs updating. According to this notation, $B_{it+s}^{(t+1)} - B_{it+s}^{(t)}$ represents the change (or update) from period t to period $t + 1$ in the beliefs that firm i has about behavior at period $t + s$.

Given a firm's beliefs at period t, $\mathbf{B}_i^{(t)} = \{B_{i,t+s}^{(t)} : s \geq 0\}$, its best response at period t is the solution of a single-agent dynamic programming (DP) problem. We can represent this DP problem

using Bellman's principle. Let $V_{it}^{\mathbf{B}_i^{(t)}}(\mathbf{x}_t, \varepsilon_{it})$ be the value function. Then,

$$V_{it}^{\mathbf{B}_i^{(t)}}(\mathbf{x}_t, \varepsilon_{it}) = \max_{a_{it} \in \mathcal{A}} \left\{ \int_{\boldsymbol{a}_{-it}} \left[\pi_{it}(a_{it}, \boldsymbol{a}_{-it}, \varepsilon_{it}, \mathbf{x}_t) + v_{it}^{\mathbf{B}_i^{(t)}}(a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t) \right] B_{it}^{(t)}(\boldsymbol{a}_{-it}|\mathbf{x}_t) \, d\boldsymbol{a}_{-it} \right\} \quad (19)$$

where $v_{it}^{\mathbf{B}_i^{(t)}}(a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t)$ is the continuation value:

$$\beta \int V_{it+1}^{\mathbf{B}_i^{(t)}}(\mathbf{x}_{t+1}, \varepsilon_{it+1}) \, f_{xt}(\mathbf{x}_{t+1}|a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t) \, d\mathbf{x}_{t+1} \, dF_i(\varepsilon_{it+1}|\mathbf{x}_t) \quad (20)$$

Given its beliefs, a firm chooses its strategy function $\sigma_{it}(\varepsilon_{it}, \mathbf{x}_t, \mathbf{B}_i^{(t)})$ to maximize its expected value:

$$\sigma_{it}(\varepsilon_{it}, \mathbf{x}_t, \mathbf{B}_i^{(t)}) = \arg \max_{a_{it} \in \mathcal{A}} \left\{ \int_{\boldsymbol{a}_{-it}} \left[\pi_{it}(a_{it}, \boldsymbol{a}_{-it}, \varepsilon_{it}, \mathbf{x}_t) + v_{it}^{\mathbf{B}_i^{(t)}}(a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t) \right] B_{it}^{(t)}(\boldsymbol{a}_{-it}|\mathbf{x}_t) \, d\boldsymbol{a}_{-it} \right\}$$
$$(21)$$

As in the static game, it is convenient to represent a firm's strategy as a cumulative distribution function. Let $P_{it}(a_{it}|\mathbf{x}_t)$ be the cumulative distribution of the choice variable a_{it} conditional on $\mathbf{x}_t$ that we denote as the *cumulative choice probability function*. For any value $a^0 \in \mathcal{A}$,

$$P_{it}(a^0|\mathbf{x}_t) \equiv \int 1\left\{ \sigma_{it}(\varepsilon_{it}, \mathbf{x}_t, \mathbf{B}_i^{(t)}) \leq a^0 \right\} \, dF_i(\varepsilon_{it}|\mathbf{x}_t)$$

EXAMPLE 5 (Continuous decision variable). Consider a dynamic game where the decision variable a_{it} is continuous. Suppose that the marginal profit function is additive in ε_{it}: $\Delta\pi_{it}(a_{it}, \boldsymbol{a}_{-it}, \varepsilon_{it}, \mathbf{x}_t) = \Delta\pi_{it}^*(a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t) - \varepsilon_{it}$. In this model, the cumulative choice probability function satisfies the following equation:

$$P_{it}(a^0|\mathbf{x}_t) = F_i\left(\int_{\boldsymbol{a}_{-it}} \left[\Delta\pi_{it}^*(a^0, \boldsymbol{a}_{-it}, \mathbf{x}_t) + \Delta v_{it}^{\mathbf{B}_i^{(t)}}(a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t) \right] B_{it}^{(t)}(\boldsymbol{a}_{-it}|\mathbf{x}_t) \, d\boldsymbol{a}_{-it} \mid \mathbf{x} \right) \quad (22)$$

where $\Delta v_{it}^{\mathbf{B}_i^{(t)}}(a_{it}, \boldsymbol{a}_{-it}, \mathbf{x}_t)$ is the marginal continuation value with respect to a_{it}. ∎

EXAMPLE 6 (Binary decision variable). Consider a dynamic game of market entry and exit with $a_{it} \in \{0, 1\}$. Suppose that the marginal profit function is additive in ε_{it}: $\Delta\pi_{it}(1, \boldsymbol{a}_{-it}, \varepsilon_{it}, \mathbf{x}_t) = \Delta\pi_{it}^*(1, \boldsymbol{a}_{-it}, \mathbf{x}_t) - \varepsilon_{it}$. In this model, the choice probability function satisfies the following equation:

$$P_{it}(0|\mathbf{x}_t) = 1 - F_i\left(\sum_{\boldsymbol{a}_{-it}} \left[\Delta\pi_{it}^*(1, \boldsymbol{a}_{-it}, \mathbf{x}_t) + \Delta v_{it}^{\mathbf{B}_i^{(t)}}(1, \boldsymbol{a}_{-it}, \mathbf{x}_t) \right] B_{it}^{(t)}(\boldsymbol{a}_{-it}|\mathbf{x}_t) \mid \mathbf{x} \right) \quad (23)$$

where $\Delta v_{it}^{\mathbf{B}_i^{(t)}}\left(1, \boldsymbol{a}_{-it}, \mathbf{x}_t\right)$ is the marginal continuation value $v_{it}^{\mathbf{B}_i^{(t)}}\left(1, \boldsymbol{a}_{-it}, \mathbf{x}_t\right) - v_{it}^{\mathbf{B}_i^{(t)}}\left(0, \boldsymbol{a}_{-it}, \mathbf{x}_t\right)$. $\blacksquare$

In this framework, the sequence of beliefs $\mathbf{B}_i^{(t)} = \left\{ B_{i,t+s}^{(t)} : s \geq 0 \right\}$ is completely unrestricted. This framework contains as particular cases most solution concepts in dynamics games of competition with incomplete information. We present here several cases.

(a) **Markov Perfect Equilibrium (MPE).** This is the most commonly used solution concept in applications of dynamic games in empirical IO. Here we consider a version of MPE that includes the possibility of non-stationary MPE when the time horizon T is finite or/and the primitive functions π_{it} and f_{xt} vary over time.

A MPE can be described as N sequences of cumulative choice probability functions – $\{P_{it}(a_{it}|\mathbf{x}_t) : i = 1, 2, ..., N; t \geq 1\}$ – satisfying the following conditions: (i) [best responses] $P_{it}(a_{it}|\mathbf{x}_t)$ satisfies the best response condition given beliefs $\mathbf{B}_i^{(t)}$; and (ii) [rational beliefs] beliefs $\mathbf{B}_i^{(t)}$ are equal to the actual probability distribution of the choices of the other firms: for any $t \geq 1$, $s \geq 0$, $\boldsymbol{a}_{-i} \in \mathcal{A}^{N-1}$, and $\mathbf{x} \in \mathcal{X}$,

$$B_{i,t+s}^{(t)}(\boldsymbol{a}_{-i}|\mathbf{x}) = \prod_{j \neq i} \Delta P_{j,t+s}\left(a_j|\mathbf{x}\right) \tag{24}$$

where $\Delta P_{jt}\left(a_j|\mathbf{x}\right)$ is the density probability function associated to the cumulative distribution $P_{jt}(a_j|\mathbf{x})$.

(b) **Dynamic equilibrium with Learning.** Bayesian, Adaptive, or other forms of learning. All these equilibrium concepts impose some restrictions on beliefs: both on the heterogeneity of firms' beliefs and on the evolution of beliefs over time. They are restricted versions of this general model.

4   Identification

4.1   Data

This section presents results on the identification of beliefs in the previous general framework and under different types of data which are common in empirical applications in IO. I distinguish three possible scenarios for the data available to the researcher.

(a) Only firms' choice data. The researcher has a sample of M local markets, indexed by m,

where she observes firms' actions and state variables:

$$\text{Data} = \{a_{imt}, \ \mathbf{x}_{mt} : i = 1, 2, ..., N; \ t = 1, 2, ..., T^{data}\}$$

This is typically described as *firms' choice data*. In empirical applications of market entry models, it is often the case that the researcher has only choice data – e.g., firms' entry/exit decision – and there is no direct information on firms' revenues or costs.

(b) Choice data + revenue function. In addition to data on firms' choices, the researcher may have data on some components of the profit function. In many IO applications, the researcher observes prices and quantities and can estimate the demand system. Given the demand system, the researcher knows the revenue function, and therefore the marginal revenue function.

(c) Choice data + revenue function + cost function. Data on firms' marginal costs is rare but it is sometimes available (Hortaçsu and Puller, 2008; Hortaçsu et al., 2019). Marginal costs can be also obtained from the estimation of a production function if the dataset contains information on firms' output and inputs, and input prices.

To incorporate in our framework the data that the researcher has on the revenue or cost functions, we distinguish these two components in the profit function. A firm's profit is equal to revenue minus cost: $\pi_i = r_i - c_i$. Accordingly, the marginal profit is equal to the marginal revenue minus the marginal cost: $\Delta \pi_i = \Delta r_i - \Delta c_i$. As explained above, this marginal revenue and marginal cost should be interpreted in a broad sense because they depend on the particular decision variable of the model that can be continuous – e.g., quantity, price, investment – or discrete – e.g., entry, number of products.

For the identification analysis below, I consider that the researcher has a random sample with infinite markets: $M \to \infty$. This is quite standard in the literature on identification. Given this infinite sample, the cumulative choice probability functions $P_i(a^0|\mathbf{x}^0)$ are identify at every value $(a^0, \mathbf{x}^0)$ in the support $\mathcal{A} \times \mathcal{X}$ and for every firm i. More precisely, we have that

$$P_i(a^0|\mathbf{x}^0) = \mathbb{E}\left(1\left\{a_{im} \leq a^0\right\} \mid \mathbf{x}_m = \mathbf{x}^0\right) \tag{25}$$

and the expectation $\mathbb{E}\left(1\left\{a_{im} \leq a^0\right\} \mid \mathbf{x}_m = \mathbf{x}^0\right)$ is identified from our sample. For the rest of this section, I treat P_i as a known function.

4.2 The identification problem

Consider the static game in section 3.1. Proposition 1 establishes that, for any $a^0 \in \mathcal{A}$:

$$P_i(a^0|\mathbf{x}) = F_i\left(e_i^*(a^0, \mathbf{x}, B_i) \mid \mathbf{x}\right) \tag{26}$$

where $e_i^*(a^0, \mathbf{x}, B_i)$ is the function that results from solving for ε_i in equation $\Delta \pi_i^e(a^0, \varepsilon_i, x; B_i) = 0$. For notational simplicity, I concentrate in the case where the private information ε_i enters additively in the marginal profit:

$$\Delta \pi_i(a_i, \boldsymbol{a}_{-i}, \varepsilon_i, \mathbf{x}) = \Delta r_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) - \Delta c_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) - \varepsilon_i \tag{27}$$

Under this condition, we have that function $e_i^*(a^0, \mathbf{x}, B_i)$ is equal to the expected marginal profit: $e_i^*(a^0, \mathbf{x}, B_i) = \int [\Delta r_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) - \Delta c_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x})] \, B_i(\boldsymbol{a}_{-i}|\mathbf{x}) \, d\boldsymbol{a}_{-i}$. Therefore, the best response conditions of the model can be written as:

$$P_i(a^0|\mathbf{x}) = F_i\left(\Delta r_i^e(a^0, \mathbf{x}, B_i) - \Delta c_i^e(a^0, \mathbf{x}, B_i) \mid \mathbf{x}\right) \tag{28}$$

where $\Delta r_i^e(a^0, \mathbf{x}, B_i) \equiv \int \Delta r_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i}|\mathbf{x}) \, d\boldsymbol{a}_{-i}$ and $\Delta c_i^e(a^0, \mathbf{x}, B_i) \equiv \int \Delta c_i(a_i, \boldsymbol{a}_{-i}, \mathbf{x}) \, B_i(\boldsymbol{a}_{-i}|\mathbf{x}) \, d\boldsymbol{a}_{-i}$ are the (subjective) expected marginal revenue and expected marginal cost, respectively.

Here we are particularly interested in the identification of the belief functions $\{B_i(\boldsymbol{a}_{-i}|\mathbf{x})\}$. I consider identification results that do not rely on parametric assumptions, neither on beliefs nor on the marginal revenue and cost functions, or the distributions of the private information, F_i.

Equation (28) summarizes all the restrictions that the model imposes on the distribution function P_i. The left-hand-side of this equation – the distribution P_i – is known to the researcher. The right-hand-side depends on the model primitives – the structural functions F_i, Δr_i, and Δc_i – and on beliefs B_i. The identification problem consists in obtaining firms' beliefs, marginal revenue, marginal cost, and the distribution of private information given the available data and the restrictions in (28).

It is clear that the model is strongly under-identified. While the number of restrictions – the dimension of the distribution P_i – is $|\mathcal{A}|^{N-1} \, |\mathcal{X}|$, we have that only the dimension of the beliefs function B_i is $|\mathcal{A}|^{N-1} \, |\mathcal{X}|$ which is obviously larger than $|\mathcal{A}| \, |\mathcal{X}|$. When F_i, Δr_i, and Δc_i are unknown to the researcher, the under-identification is stronger. Despite the under-identification of the model, I show below that it possible to identify a function that only depends of beliefs. The identification of this "beliefs object" can be used to test the validity of different equilibrium concepts and restrictions on beliefs.

16

For the sake of simplicity, I first illustrate the identification results in a simple model of competition: a binary choice game with two firms. Furthermore, I assume that the probability distribution F_i is known to the researcher. Then, in section 4.4, I extend the identification results to: (i) more than two players; (ii) multinomial and continuous choice models; (iii) nonparametric specification F_i; and (iv) dynamic games.

4.3 Two-firms binary choice game

Consider a binary choice game of price competition between two firms: $a_i = 0$ represents the choice of the low price (promotion price) and $a_i = 1$ represents the choice of high price (regular price). Let $q_i = d_i(a_i, a_{-i}, \mathbf{x})$ be the demand function for the product of firm i – i.e., quantity as a function of prices; and let $C_i(q_i, \mathbf{x})$ be the cost of a firm as a function of its own output. Therefore, using our notation, the revenue function is $r_i(a_i, a_{-i}, \mathbf{x}) = a_i \, d_i(a_i, a_{-i}, \mathbf{x})$, and the cost function is $r_i(a_i, a_{-i}, \mathbf{x}) = C_i(d_i(a_i, a_{-i}, \mathbf{x}), \mathbf{x}).$[2]

Let $P_i(0|\mathbf{x})$ – or in short $P_i(\mathbf{x})$ – be the probability that firm i chooses the low price. And let $B_i(0|\mathbf{x})$ – or in short $B_i(\mathbf{x})$ – be this firm's belief about the probability that the competitor chooses the low price. The marginal profit function is $\Delta\pi_i(a_{-i}, \mathbf{x}) \equiv \pi_i(1, a_{-i}, \mathbf{x}) - \pi_i(0, a_{-i}, \mathbf{x})$, that is, the difference between the profit with high price and with low price. Marginal profit is equal to marginal revenue minus marginal cost: $\Delta\pi_i(a_{-i}, \mathbf{x}) = \Delta r_i(a_{-i}, \mathbf{x}) - \Delta c_i(a_{-i}, \mathbf{x})$.

The model can be described in terms of the best response equation:

$$P_i(\mathbf{x}) = F_i\left(\Delta\pi_i(0, \mathbf{x}) + B_i(\mathbf{x})\,[\Delta\pi_i(1, \mathbf{x}) - \Delta\pi_i(0, \mathbf{x})]\right) \tag{29}$$

Define the quantile function $Q_i(\mathbf{x}) \equiv F_i^{-1}(P_i(\mathbf{x}))$. Under the assumption that the distribution of the private information, F_i, is known to the researcher, the quantile function $Q_i(\mathbf{x})$ is also known, and we can represent the restrictions of the model using the following equation:

$$Q_i(\mathbf{x}) = \Delta\pi_i(0, \mathbf{x}) + B_i(\mathbf{x})\,[\Delta\pi_i(1, \mathbf{x}) - \Delta\pi_i(0, \mathbf{x})] \tag{30}$$

Given that $Q_i(\mathbf{x})$ is known, we are interested in the identification of marginal profits $\Delta\pi_i(0, \mathbf{x})$ and $\Delta\pi_i(1, \mathbf{x})$, and the beliefs function $B_i(\mathbf{x})$. We are particularly interested in the identification of the beliefs function $B_i(\mathbf{x})$, or at least on the identification of an object or parameter that only depends on this belief function.

---

[2] Even if a firm's cost depends only on its own output, the cost as a function of prices depends both on the own price and competitors' prices. This is simply because the quantity produced and sold by a firm depends on all the prices. In contrast, in a Cournot game where a_i represents a firm's output, the cost function $c_i(a_i, a_{-i}, \mathbf{x})$ does not depend on a_{-i}.

Without further restrictions, it is clear that the model is under-identified. More specifically, the order condition for identification does not hold: for each value of $\mathbf{x}$, there is only one restriction – i.e., one value of $Q_i(\mathbf{x})$ – but three unknowns, $\Delta\pi_i(0, \mathbf{x})$, $\Delta\pi_i(1, \mathbf{x})$, and $B_i(\mathbf{x})$. I now describe how it is possible to make some progress on identification.

4.3.1   Identification with revenue and cost data

Suppose that researcher knows the revenue function and the cost function. For instance, the dataset includes information on prices and quantities of inputs and outputs – as well as exogenous variables – that can be used to identify demand and cost functions. This implies that the marginal profits $\Delta\pi_i(0, \mathbf{x})$ and $\Delta\pi_i(1, \mathbf{x})$ are known to the researcher. Therefore, under the condition that $\Delta\pi_i(1, \mathbf{x}) - \Delta\pi_i(0, \mathbf{x}) \neq 0$, the beliefs function is fully identified:

$$B_i(\mathbf{x}) = \frac{Q_i(\mathbf{x}) - \Delta\pi_i(0, \mathbf{x})}{\Delta\pi_i(1, \mathbf{x}) - \Delta\pi_i(0, \mathbf{x})} \tag{31}$$

The identification condition $\Delta\pi_i(1, \mathbf{x}) - \Delta\pi_i(0, \mathbf{x}) \neq 0$ is quite intuitive: a firm's observed behavior reveals information about the firm's beliefs only if beliefs have an effect on behavior, and this is the case only if other firms' actions affect the firm's profit, i.e., only if $\Delta\pi_i(1, \mathbf{x}) - \Delta\pi_i(0, \mathbf{x}) \neq 0$.

Given the identification of firms' beliefs, the researcher can test the validity of different types of restrictions on beliefs.

(a) Testing for unbiased beliefs. We say that firm i has unbiased beliefs about the behavior of the other firm if $B_i(\mathbf{x}) - P_{-i}(\mathbf{x}) = 0$ for every value of $\mathbf{x}$. Given the identification of $B_i(\mathbf{x})$ and that $P_{-i}(\mathbf{x})$ is known to the researcher, we can test the null hypothesis of a firm's unbiased beliefs.

(b) Testing for BNE. The concept of BNE imposes the restrictions that all the firms play best responses and have unbiased beliefs. Therefore, given the identification of firms' belief functions $B_1(\mathbf{x})$ and $B_2(\mathbf{x})$ – which come from the best response equations – testing the null hypothesis of BNE is equivalent to test the joint restrictions $B_1(\mathbf{x}) - P_2(\mathbf{x}) = 0$ and $B_2(\mathbf{x}) - P_1(\mathbf{x}) = 0$ for every value of $\mathbf{x}$.

(c) Testing for Rationalizabilitty. Given that the researcher knows the profit functions, she can construct the set of rationalizable beliefs, and then test if the identified beliefs – $B_1(\mathbf{x})$ and $B_2(\mathbf{x})$ – belong to this set. To construct the set of rationalizable beliefs we can use a simple iterative procedure as in Aradillas-Lopez and Tamer (2008). This iterative procedure exploits the property

that the best response probability function $F_i(\Delta\pi_i(0,\mathbf{x}) + B_i(\mathbf{x})[\Delta\pi_i(1,\mathbf{x}) - \Delta\pi_i(0,\mathbf{x})])$ is strictly monotonic in the beliefs function $B_i(\mathbf{x})$. Suppose that $\Delta\pi_i(1,\mathbf{x}) - \Delta\pi_i(0,\mathbf{x}) > 0$ such that in this game of price competition there is strategic complementarity between the prices of the two firms. At iteration k, the set of level-k rationalizable beliefs for firm 1 is $[L_1^{(k)}(\mathbf{x}), U_1^{(1)}(\mathbf{x})]$ with

$$
\begin{aligned}
L_1^{(k)}(\mathbf{x}) &= F_2(\Delta\pi_2(0,\mathbf{x}) + L_2^{(k-1)}(\mathbf{x})[\Delta\pi_2(1,\mathbf{x}) - \Delta\pi_2(0,\mathbf{x})]) \\
U_1^{(k)}(\mathbf{x}) &= F_2(\Delta\pi_2(0,\mathbf{x}) + U_2^{(k-1)}(\mathbf{x})[\Delta\pi_2(1,\mathbf{x}) - \Delta\pi_2(0,\mathbf{x})])
\end{aligned}
\tag{32}
$$

And we have the symmetric expression for the set of level-k rationalizable beliefs for firm 2.

4.3.2 Identification with revenue but not cost data

Suppose that the researcher has data that identifies the demand system and therefore the revenue function $r_i(a_i, a_{-i}, \mathbf{x})$. The cost function is unknown. We have the best response equation:

$$
Q_i(\mathbf{x}) = \Delta r_i(0,\mathbf{x}) - \Delta c_i(0,\mathbf{x}) + B_i(\mathbf{x})[\Delta r_i(1,\mathbf{x}) - \Delta r_i(0,\mathbf{x}) + \Delta c_i(1,\mathbf{x}) - \Delta c_i(0,\mathbf{x})]
\tag{33}
$$

where $Q_i(\mathbf{x})$, $\Delta r_i(0,\mathbf{x})$, and $\Delta r_i(1,\mathbf{x})$ are known, and $B_i(\mathbf{x})$, $\Delta c_i(1,\mathbf{x})$, and $\Delta c_i(0,\mathbf{x})$ are unknown. It is clear that this restriction cannot identify beliefs and cost functions. For any possible value of $B_i(\mathbf{x})$, there exists a value of $\Delta c_i(1,\mathbf{x}) - \Delta c_i(0,\mathbf{x})$ such that the best response equation holds.

This identification problem is closely related to the identification of collusion – or for that matter, any form of competition – when the researcher does not have information on marginal costs (Breshnahan, 1982). Almost any observed behavior can be justified as one with "non-collusive beliefs" if we select the appropriate marginal cost function.

Firms' beliefs and conjectural variations   In an influential paper, Breshnahan (1982) studies the identification of the *form (or nature) of competition* in a model with complete information. In a complete information game, the nature of competition can be described as a conjectural variation (CV) parameter. This CV parameter has similarities with our beliefs function, but there are also substantial differences between them. Our beliefs function is an endogenous object that varies with all the exogenous characteristics in the vector $\mathbf{x}$ affecting demand and costs. CV parameters are typically interpreted as exogenously given and do not vary when demand or costs change. As I explain below, this has important implications of the identification of beliefs relative to the identification of CV parameters.

The best response equation in Bresnahan (1982) is similar as equation (33) but replacing the beliefs function $B_i(\mathbf{x})$ with a parameter CV_i that is assumed invariant with $\mathbf{x}$. After the identification of demand equation and the marginal revenue function, Bresnahan proposes an exclusion

19

restriction that implies the identification of the parameter CV_i. I first describe this identification result using our notation, and then I show this assumption cannot provide identification of beliefs in our model.

Suppose that the vector of exogenous variables $\mathbf{x}$ has two components $(\widetilde{\mathbf{x}}, z)$ where z is a variable that satisfies two conditions: (i) it affects the marginal revenue function, or more precisely the function $\Delta r_i(1, \mathbf{x}) - \Delta r_i(0, \mathbf{x})$ – it "rotates" the demand curve; and (ii) it does not enter in the marginal cost function.

Consider the best response equation (33) but where the beliefs function $B_i(\mathbf{x})$ is replaced with the parameter CV_i. Let z^1 and z^2 be two different values of the special variable that "rotates" the demand curve. Consider the best response equation evaluated at two different points, $(\widetilde{\mathbf{x}}, z^1)$ and $(\widetilde{\mathbf{x}}, z^2)$, and obtain the difference between these two equations. We get:

$$
\begin{aligned}
Q_i(\widetilde{\mathbf{x}}, z^1) - Q_i(\widetilde{\mathbf{x}}, z^2) \; = \;\; & \Delta r_i(0, \widetilde{\mathbf{x}}, z^1) - \Delta r_i(0, \widetilde{\mathbf{x}}, z^2) \\[2mm]
+ \;\; & CV_i \; \left[\Delta r_i(1, \widetilde{\mathbf{x}}, z^1) - \Delta r_i(0, \widetilde{\mathbf{x}}, z^1) - \Delta r_i(1, \widetilde{\mathbf{x}}, z^2) + \Delta r_i(0, \widetilde{\mathbf{x}}, z^2) \right]
\end{aligned}
\tag{34}
$$

Everything in this equation except parameter CV_i is known to the researcher. Furthermore, the identification assumption (ii) above implies that $\Delta r_i(1, \widetilde{\mathbf{x}}, z^1) - \Delta r_i(0, \widetilde{\mathbf{x}}, z^1) - \Delta r_i(1, \widetilde{\mathbf{x}}, z^2) + \Delta r_i(0, \widetilde{\mathbf{x}}, z^2)$ is different to zero. Therefore, we can solve for CV_i to identify this parameter.

However, this exclusion restriction does not work for the identification of beliefs. In general, the beliefs function $B_i(\mathbf{x})$ depends on all the exogenous variables affecting demand or costs. Therefore, under the identification assumptions (i) and (ii) above, we have that right-hand-side of equation (34) becomes:

$$
\begin{aligned}
& \Delta r_i(0, \widetilde{\mathbf{x}}, z^1) - \Delta r_i(0, \widetilde{\mathbf{x}}, z^2) + \left[B_i(\widetilde{\mathbf{x}}, z^1) - B_i(\widetilde{\mathbf{x}}, z^2) \right] \; \left[\Delta c_i(1, \widetilde{\mathbf{x}}) - \Delta c_i(0, \widetilde{\mathbf{x}}) \right] \\[2mm]
& + B_i(\widetilde{\mathbf{x}}, z^1) \; \left[\Delta r_i(1, \widetilde{\mathbf{x}}, z^1) - \Delta r_i(0, \widetilde{\mathbf{x}}, z^1) \right] - B_i(\widetilde{\mathbf{x}}, z^2) \; \left[\Delta r_i(1, \widetilde{\mathbf{x}}, z^2) - \Delta r_i(0, \widetilde{\mathbf{x}}, z^2) \right]
\end{aligned}
\tag{35}
$$

This expression depends both on beliefs and costs and it cannot be used to separately identify one from the other.

An identifying restriction Suppose that the vector $\mathbf{x}$ contains a firm-specific variable that affects the marginal cost of a firm but not the marginal cost of its competitors. For instance, input prices – wages, prices of intermediate inputs – can have firm specific variation because long-term contracts, bargaining, internal labor markets, etc. Formally, the vector $\mathbf{x}$ has three components $(\widetilde{\mathbf{x}}, z_i, z_{-i})$ where $\widetilde{\mathbf{x}}$ can affect the marginal revenue and marginal costs of the two firms in an unrestricted way, and the variables z_i and z_{-i} satisfy the following conditions.

(A) Firms' marginal revenues do not depend on (z_i, z_{-i}): $\Delta r_i(a_{-i}, \mathbf{x}) = \Delta r_i(a_{-i}, \widetilde{\mathbf{x}})$.

(B) The marginal cost of firm i depends on z_i but not on z_{-i}: $\Delta c_i(a_{-i}, \mathbf{x}) = \Delta c_i(a_{-i}, \widetilde{\mathbf{x}}, z_i)$.

Under conditions (A) and (B), there is identification of an object that depends only on beliefs. The term $Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1)$ captures the change in the behavior of firm i when z_{-i} changes from z_{-i}^1 to z_{-i}^2: that is, the change in the probability that firm i charges a low price when the competitor's wage rate changes. Since variable z_{-i} does not affect firm i's marginal revenue or marginal cost, it is clear that the change the observed change in the pricing behavior in firm i should be because a change in its beliefs. It turns out that this shift identifies a function that depends only on beliefs.

The difference between the best-response equation at points $(\widetilde{\mathbf{x}}, z_i, z_{-i}^1)$ and $(\widetilde{\mathbf{x}}, z_i, z_{-i}^2)$ is:

$$Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2) = \left[B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2) \right] \left[\Delta \pi_i(1, \widetilde{\mathbf{x}}, z_i) - \Delta \pi_i(0, \widetilde{\mathbf{x}}, z_i) \right] \quad (36)$$

This difference is not sufficient to identify the beliefs parameter $B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2)$. The reason, is that $\Delta \pi_i(1, \widetilde{\mathbf{x}}, z_i) - \Delta \pi_i(0, \widetilde{\mathbf{x}}, z_i)$ depends on unknown marginal costs through the term $\Delta c_i(1, \widetilde{\mathbf{x}}, z_i) - \Delta c_i(0, \widetilde{\mathbf{x}}, z_i)$. However, we can also obtain the difference between the best-response equation at points $(\widetilde{\mathbf{x}}, z_i, z_{-i}^1)$ and $(\widetilde{\mathbf{x}}, z_i, z_{-i}^3)$ to get:

$$Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^3) = \left[B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^3) \right] \left[\Delta \pi_i(1, \widetilde{\mathbf{x}}, z_i) - \Delta \pi_i(0, \widetilde{\mathbf{x}}, z_i) \right] \quad (37)$$

Note that the term $\Delta \pi_i(1, \widetilde{\mathbf{x}}, z_i) - \Delta \pi_i(0, \widetilde{\mathbf{x}}, z_i)$ is common between these equations. Therefore, we can cancel this unknown common term by obtaining the ratio between these two equations, and identify the following beliefs parameter:

$$\frac{B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2)}{B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^3)} = \frac{Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2)}{Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^3)} \quad (38)$$

Equation (38) shows that the observed variation in the pricing behavior of firm i – when the competitor's input prices change – reveals information about this firm's beliefs. We can separate beliefs from the primitives in the profit function.

In some models, the cost function of a firm does not depend on the action of other firms. For instance, this is the case in Cournot models of quantity competition or in the entry games because, in these models, the cost function c_i is a "pure" cost function and not the composition of the true cost function and the demand function. In these models, we have that $\Delta c_i(1, \widetilde{\mathbf{x}}, z_i) - \Delta c_i(0, \widetilde{\mathbf{x}}, z_i) = 0$

such that $\Delta\pi_i(1,\widetilde{\mathbf{x}},z_i) - \Delta\pi_i(0,\widetilde{\mathbf{x}},z_i) = \Delta r_i(1,\widetilde{\mathbf{x}}) - \Delta r_i(0,\widetilde{\mathbf{x}})$ that is known to the researcher. Therefore, we can identify the following beliefs parameter:

$$B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - B_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2) = \frac{Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2)}{\Delta r_i(1,\widetilde{\mathbf{x}}) - \Delta r_i(0,\widetilde{\mathbf{x}})} \tag{39}$$

Given the identification of these beliefs objects, we can implement tests for the null hypotheses of unbiased beliefs and BNE in a similar way as I have described above at the end of section 4.3.1.

Define the following functions: $q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1, z_{-i}^2, z_{-i}^3) \equiv [Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^2)] / [Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - Q_i(\widetilde{\mathbf{x}}, z_i, z_{-i}^3)]$; and $p_{-i}(\widetilde{\mathbf{x}}, z_i, z_{-i}^1, z_{-i}^2, z_{-i}^3) \equiv [P_{-i}(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - P_{-i}(\widetilde{\mathbf{x}}, z_i, z_{-i}^2)] / [P_{-i}(\widetilde{\mathbf{x}}, z_i, z_{-i}^1) - P_{-i}(\widetilde{\mathbf{x}}, z_i, z_{-i}^3)]$. To shorten the notation, we use q_i and p_{-i}, respectively, to represent these functions.

(a) Testing for unbiased beliefs. We can test the null hypothesis of firm i's unbiased beliefs by testing the restrictions: $q_i - p_{-i}$.

(b) Testing for BNE. We can test the null hypothesis of BNE by testing the restrictions: $q_1 - p_2 = 0$ and $q_2 - p_1 = 0$.

4.3.3 Identification using only firms' choice data

The previous exclusion restriction can be applied to the identification of beliefs also when the researcher has not identified the revenue function. Now, we need that the variable that shifts the marginal cost of the competitor has at least three points of support. We can derive equation (38) exactly in the same way as described above. Similarly, we can use the identified beliefs parameters to test the null hypotheses of unbiased beliefs and BNE.

4.4 Extensions

4.4.1 Identification of beliefs with nonparametric distribution of private information

When the decision variable a_i is continuous or discrete but ordered variable with more than two values, it is possible to obtain identification of a function of beliefs that does not depend on the assumption that the distribution F_i is known to the researcher.

Remember that the general model can be described in terms of the following restrictions: for any value $a^0 > 0$:

$$F_i^{-1}\left[\Pr(a_i \geq a^0 \mid \mathbf{x})\right] = r_i^{B_i}(a^0, \mathbf{x}) - c_i(a^0, \mathbf{x}) \tag{40}$$

Let $\mathbf{z}^1_{-i}$ and $\mathbf{z}^2_{-i}$ be two values for the shifter of the opponent's marginal cost. Suppose that there are two values for the price a_i, say a^1 and a^2 such that:

$$\Pr(a_i \geq a^1 \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}^1_{-i}) = \Pr(a_i \geq a^2 \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}^2_{-i}) \tag{41}$$

**** EXPLAIN THE CONDITIONS UNDER WHICH THESE VALUES a^1 and a^2 ALWAYS EXIST ***

This implies that $F_i^{-1}\left[\Pr(a_i \geq a^1 \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}^1_{-i})\right] = F_i^{-1}\left[\Pr(a_i \geq a^2 \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}^2_{-i})\right]$. Therefore, we can take the difference between the best responses for $\Pr(a_i \geq a^1 \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}^1_{-i})$ and for $\Pr(a_i \geq a^2 \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}^2_{-i})$ and obtain an expression that does not depend on the distribution of the private information.

$$r_i^{B_i}(a^1, \mathbf{w}, \mathbf{z}_i, \mathbf{z}^1_{-i}) - r_i^{B_i}(a^2, \mathbf{w}, \mathbf{z}_i, \mathbf{z}^2_{-i}) - c_i(a^1, \mathbf{z}_i) + c_i(a^2, \mathbf{z}_i) = 0 \tag{42}$$

Or,

$$r_i(a^1_i, 0, \mathbf{w}) + \sum_{a_{-i} > 0} B_i(a_{-i} \mid \mathbf{x}^1)\left[r_i(a^1, a_{-i}, \mathbf{w}) - r_i(a^1, 0, \mathbf{w})\right] - c_i(a^1, \mathbf{z}_i)$$

$$= r_i(a^2, 0, \mathbf{w}) + \sum_{a_{-i} > 0} B_i(a_{-i} \mid \mathbf{x}^2)\left[r_i(a^2, a_{-i}, \mathbf{w}) - r_i(a^2, 0, \mathbf{w})\right] - c_i(a^2, \mathbf{z}_i)$$

*** EXPLAIN DOUBLE DIFFERENCE: ....

4.4.2   More than two players

Let's continue with the binary choice game, but now consider that there are $N > 2$ firms. The best response function can be written as:

$$Q_i(\mathbf{x}) = r_i(\mathbf{0}, \mathbf{x}) + \sum_{\boldsymbol{a}_{-i} \neq \mathbf{0}} B_i(\boldsymbol{a}_{-i} \mid \mathbf{x})\ [r_i(\boldsymbol{a}_{-i}, \mathbf{x}) - r_i(\mathbf{0}, \mathbf{x})] - c_i(\mathbf{x})$$

For instance, with $N = 3$ players:

$$
\begin{aligned}
Q_i(\mathbf{x}) \ = \ & -c_i(\mathbf{x}) + r_i([0,0], \mathbf{x}) + B_i([0,1] \mid \mathbf{x})\ [r_i([0,1], \mathbf{x}) - r_i([0,0], \mathbf{x})] \\
+ \ & B_i([1,0] \mid \mathbf{x})\ [r_i([1,0], \mathbf{x}) - r_i([0,0], \mathbf{x})] + B_i([1,1] \mid \mathbf{x})\ [r_i([1,1], \mathbf{x}) - r_i([0,0], \mathbf{x})]
\end{aligned}
$$

Even if the researcher knows the marginal revenue function and the cost function, there may be infinite values of $B_i([0,0] \mid \mathbf{x})$, $B_i([0,1] \mid \mathbf{x})$, and $B_i([1,1] \mid \mathbf{x})$ that can rationalize the observed behavior $Q_i(\mathbf{x})$.

However, the exclusion restriction of a firm-specific cost shifter still implies identification of beliefs even when the cost function or the revenue function are not known to the researcher. The identification result that is simpler to describe is the one the cost function is not known but the revenue function is known. The space of the possible actions $\boldsymbol{a}_{-i} \neq 0$ has $2^N - 1$ values. Consider $2^N - 1$ different values for $\mathbf{z}_{-i}$, this is feasible even in the cost shifter of each firm can take only

two values. Let's consider the system of best response functions associated with this $2^N - 1$ values. In vector form:

$$\mathbf{Q}_i(\mathbf{w}, \mathbf{z}_i, .) - Q_i(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1) = diag\left\{[r_i(., \mathbf{w}) - r_i(\mathbf{0}, \mathbf{w})]\right\}\ \left[\mathbf{B}_i(. \mid \mathbf{w}, \mathbf{z}_i, .) - B_i(. \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1)\right]$$

and this implies that $\mathbf{B}_i(. \mid \mathbf{w}, \mathbf{z}_i, .)$ is identified as:

$$\mathbf{B}_i(. \mid \mathbf{w}, \mathbf{z}_i, .) - B_i(. \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1) = diag\left\{[r_i(., \mathbf{w}) - r_i(\mathbf{0}, \mathbf{w})]\right\}^{-1}\ \left[\mathbf{Q}_i(\mathbf{w}, \mathbf{z}_i, .) - Q_i(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1)\right]$$

Therefore, the exclusion restriction provides identification of beliefs $\mathbf{B}_i(. \mid \mathbf{w}, \mathbf{z}_i, .) - B_i(. \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1)$.

Without knowledge of the revenue function, it is possible to show identification of $[\mathbf{B}_i(. \mid \mathbf{w}, \mathbf{z}_i, .) - B_i(. \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}_i^1)]\ [\mathbf{B}_i(. \mid \mathbf{w}, \mathbf{z}_i, .) - B_i(. \mid \mathbf{w}, \mathbf{z}_i, \mathbf{z}_i^2)]^{-1}$.

4.4.3 Dynamic game

We have:

$$\begin{aligned}
Q_{it}(\mathbf{x}_t) &= r_{it}(0, \mathbf{x}_t) - c_{it}(\mathbf{x}_t) + \Delta V_{it}^{\mathbf{B}_{i,t+1}^{(t)}}(0, \mathbf{x}_t) \\
&+ B_{it}^{(t)}(\mathbf{x}_t)\ \left[r_{it}(1, \mathbf{x}_t) - r_{it}(0, \mathbf{x}_t) + \Delta V_{it}^{\mathbf{B}_{i,t+1}^{(t)}}(1, \mathbf{x}_t) - \Delta V_{it}^{\mathbf{B}_{i,t+1}^{(t)}}(0, \mathbf{x}_t)\right]
\end{aligned}$$

A firm-specific cost shifter that provides identification of beliefs in the dynamic game is one that does not enter in the continuation value.

Fortunately, this type of cost shifter often appears in dynamic games of oligopoly competition: the decision variable at previous period (t-1) in a model with adjustment costs.

For instance, in the model of price competition, suppose that there is cost of changing the price. A menu cost or a price adjustment cost. This implies that a firm's price at period t-1 affects its marginal cost at period t. However, given the price at period t, the price at t-1 does not have any effect on the continuation value. Therefore, we can use variation in the lagged price of the competitors(s) to differentiate out the continuation value, as well as the own cost function, and revenue function, and identify beliefs.

Let $\mathbf{x}_t = (\mathbf{w}_t, \mathbf{z}_{it}, \mathbf{z}_{-it})$ with the same properties as above, and the additional properties that the continuation values do not depend on $(\mathbf{z}_{it}, \mathbf{z}_{-it})$. Remember that the cost function may depend on different variables which are not this cost shifter. The key condition is that the competitor's cost shifter does not affect the firm's own marginal revenue, marginal cost, and continuation value.

Let $\mathbf{z}_{-i}^1$, $\mathbf{z}_{-i}^2$, and $\mathbf{z}_{-i}^2$ be three different values of this shifter. Then, we have that:

$$\frac{B_{it}^{(t)}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^2) - B_{it}^{(t)}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1)}{B_{it}^{(t)}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^3) - B_{it}^{(t)}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1)} = \frac{Q_{it}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^2) - Q_{it}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1)}{Q_{it}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^3) - Q_{it}(\mathbf{w}, \mathbf{z}_i, \mathbf{z}_{-i}^1)}$$

24

Note that we can identify the sequence of over the sample period of beliefs at period t about the opponents' contemporaneous behavior at period t. We cannot identify beliefs about the opponent's behavior in the future. However, identification of the evolution of these contemporaneous beliefs is enough for testing almost any model of learning and beliefs formation.

5   Conclusions

TBW

References

[1] Aguirregabiria, V. and A. Magesan (2017). Identification and estimation of dynamics games when players' beliefs are not in equilibrium," working paper, University of Toronto.

[2] Aguirregabiria, V., and P. Mira (2007). Sequential estimation of dynamic discrete games. *Econometrica*, 75(1), 1-53.

[3] An, Y. (2017). Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. *Journal of Econometrics*, 200, 326-343.

[4] An, Y., Y. Hu, R. Xiao (2018). Dynamic decisions under subjective expectations: A structural analysis. Manuscript. Department of Economics. Johns Hopkins University.

[5] Aradillas-Lopez, A., and E. Tamer (2008). The identification power of equilibrium in simple games. *Journal of Business & Economic Statistics*, 26, 261-283.

[6] Armantier, O., W. Bruine, G. Topa, W. van der Klaauw, and B. Zafar (2015). Inflation expectations and behavior: Do survey respondents act on their beliefs? *International Economic Review*, 56, 505-536.

[7] Armantier, O., and O. Richard (2003). Exchanges of cost information in the airline industry. *RAND Journal of Economics*, 461-477.

[8] Aryal, G., and F. Zicenko (2019). Empirical framework for Cournot oligopoly with private information. Manuscript. Department of Economics. University of Pittsburgh.

[9] Asker, J., C. Fershtman, C., J. Jeon, and A. Pakes (2016). The competitive effects of information sharing. NBER Working Paper, No. 22836. National Bureau of Economic Research.

[10] Athey, S. (2001). Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69(4), 861-889.

[11] Athey, S., and K. Bagwell (2001). Optimal Collusion with Private Information. *RAND Journal of Economics*, 32(2), 428–465.

[12] Athey, S., and P. Haile (2002). Identification of standard auction models. *Econometrica*, 70(6), 2107-2140.

[13] Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010). Estimating static models of strategic interactions. *Journal of Business & Economic Statistics*, 28, 469-482.

[14] Bergemann, D., and S. Morris (2013). Robust predictions in games with incomplete information. *Econometrica*, 81(4), 1251-1308.

[15] Bergemann, D., and S. Morris (2016). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2), 487-522.

[16] Bernheim, B. (1984). Rationalizable strategic behavior. *Econometrica*, 52, 1007-1028.

[17] Berry, S. (1992). Estimation of a Model of Entry in the Airline Industry. *Econometrica*, 60, 889-917.

[18] Berry, S. T., and P. Haile (2014). Identification in differentiated products markets using market level data. *Econometrica*, 82(5), 1749-1797.

[19] Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4), 841-890.

[20] Borkovsky, R., P. Ellickson, B. Gordon, V. Aguirregabiria, P. Gardete, P. Grieco, T. Gureckis, T. Ho, L. Mathevet, and A. Sweeting (2015). Multiplicity of equilibria and information structures in empirical games: challenges and prospects. *Marketing Letters*, 26, 115-125.

[21] Bresnahan, T. (1982). The oligopoly solution concept is identified, *Economics Letters*, 10(1-2), 87-92.

[22] Bresnahan, T., and P. Reiss (1991). Entry and competition in concentrated markets. *Journal of Political Economy*, 99(5), 977-1009.

[23] Ciliberto, F., and E. Tamer (2009). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6), 1791-1828.

[24] De Paula, A., and X. Tang (2012). Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica*, 80(1), 143-172.

[25] Doraszelski, U., G. Lewis, and A. Pakes (2018). Just starting out: Learning and equilibrium in a new market. *American Economic Review*, 108, 565-615.

[26] Doraszelski, U. and M. Satterthwaite (2010). Computable Markov-perfect industry dynamics. *Rand Journal of Economics*, 41, 215-243.

[27] Fershtman, C. and A. Pakes (2012). Dynamic games with asymmetric information: A framework for empirical work. *Quarterly Journal of Economics*, 127, 1611-1661.

[28] Gardete, P. (2016). Competing under asymmetric information: The case of dynamic random access memory manufacturing. *Management Science*, 62, 3291-3309.

[29] Genesove, D., and W. Mullin (1998). Testing static oligopoly models: conduct and cost in the sugar industry, 1890-1914. *The RAND Journal of Economics*, 29, 355-377.

[30] Goldfarb, A., and M. Xiao (2011). Who thinks about the competition? Managerial ability and strategic entry in US local telephone markets. *American Economic Review*, 101, 3130 3161.

[31] Guerre, E., I. Perrigne, and Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3), 525-574.

[32] Hortaçsu, A., and S. Puller (2008). Understanding strategic bidding in multi-unit auctions: a case study of the Texas electricity spot market. *The RAND Journal of Economics*, 39(1), 86-114.

[33] Hortaçsu, A., F. Luco, S. Puller, and D. Zhu (2019). Does strategic ability affect efficiency? Evidence from electricity markets. *American Economic Review*, 109(12), 4302-42.

[34] Igami, M. (2017). Estimating the innovator's dilemma: Structural analysis of creative destruction in the hard disk drive industry, 1981–1998. *Journal of Political Economy*, 125(3), 798-847.

[35] Manski, C. (2004). Measuring expectations. *Econometrica*, 72, 1329-1376.

[36] Magnolfi, L., and C. Roncoroni (2017). Estimation of discrete games with weak assumptions on information. Manuscript. University of Wisconsin-Madison.

[37] Morris, S. and H. Shin (2002). Measuring strategic uncertainty. Manuscript, Princeton University.

[38] Pearce, D. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52, 1029-1050.

[39] Porter, R. (1983). A study of cartel stability: the Joint Executive Committee, 1880-1886. *The Bell Journal of Economics*, 301-314.

[40] Puller, S. (2007). Pricing and firm conduct in California's deregulated electricity market. *The Review of Economics and Statistics*, 89(1), 75-87.

[41] Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics*, 37(3), 619-640.

[42] Simon, H. (1958). The role of expectations in an adaptive or behavioristic model. In M. Bowman (ed.) *Expectations, uncertainty, and business behavior*. Social Science Research Council. New York.

[43] Simon, H. (1959). Theories of decision-making in economics and behavioral science. *American Economic Review*, 49, 253-283.

[44] Sweeting, A. (2009). The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria. *The RAND Journal of Economics*, 40(4), 710-742.

[45] Xie, E. (2018). Inference in games without Nash equilibrium: An application to restaurants competition in opening hours. Manuscript. Department of Economics, University of Toronto.