

Discrete Games with Flexible Information Structures: An Application to Local Grocery Markets

Paul L. E. Grieco*

November 11, 2013

Abstract

Game-theoretic models are frequently employed to study strategic interaction between agents. Empirical research has focused on estimating payoff functions while maintaining strong assumptions regarding the information structure of the game. I show how to relax informational assumptions to enhance the credibility of empirical analysis in discrete games. I apply the method to data on the entry and exit patterns of grocery stores. The model provides useful bounds on equilibrium outcomes. In addition, the empirical analysis indicates that more restrictive informational assumptions can generate qualitatively misleading counterfactual outcomes.

Keywords: Discrete Games; Entry; Complete Information; Incomplete Information; Multiple Equilibria; Partial Identification; Grocery Retail; Supercenters.

*The Pennsylvania State University, paul.grieco@psu.edu. I gratefully acknowledge the help and support of my dissertation committee: Aviv Nevo, Rob Porter, Mark Satterthwaite and Elie Tamer. The data for this project was generously provided by Paul Ellickson. I also benefited from many conversations and discussions of earlier drafts with Guy Arie, Ivan Canay, Ulrich Doraszelski, Pierre Dubois, Jon Gemus, Gautam Gowrisankaran, Ed Green, Kei Kawai, Michael Mazzeo, Francesca Molinari, Joris Pinske, Ryan McDevitt, Mark Roberts, Katja Seim, Mauricio Valera, Chris Vickers, and Quang Vuong. I thank Ali Hortacsu and two anonymous referees for providing comments which substantially improved the article. I am also grateful for comments from seminar audiences at Johns Hopkins University, University of Wisconsin-Madison, University of Arizona, Stanford University, The Econometric Society World Congress 2010 in Shanghai, the Southern Economic Association 2010 Meetings in Atlanta, and the IIOC 2011 meetings in Boston. All remaining errors are my own.

1 Introduction

The discrete actions of firms are indicators of their expectations about profitability. Following this logic, researchers use data on firm decisions such as entry, exit, or location choice to infer the effects of market size, competition, and other factors on profitability. When multiple firms compete, their strategic interaction should be modeled as a game, and their actions as the outcome of equilibrium play. However, equilibrium outcomes depend on both firms' payoffs and firms' information when they make their decision. Hence, modeling choices about firms' information sets can affect how the data are interpreted. Previous empirical studies have typically made strong informational assumptions on firm information. In this article, I propose a more flexible approach which nests many common informational assumptions. I show how this more flexible model can be used directly for empirical work and how it can serve as a robustness check for stronger modeling assumptions.

I apply the model to study the impact of supercenters—large stores, such as Wal-Mart, that sell both food and groceries—on the profitability of rural grocery stores. Supercenters' effect on the profitability of small grocers is a matter of significant public concern. A key feature of this setting is that there is substantial unobserved heterogeneity in profitability across grocery stores; some elements are publicly known to all players (e.g., ownership of a prime retail location), and some are privately known only to that firm (e.g., managerial ability). Although it is common to abstract away from one or the other of these sources of heterogeneity, I present a model that encompasses both and allows inference on which source is dominant in the data. I also show how other methods may lead the researcher to draw conclusions on the impact of supercenters that are not supported under weaker informational assumptions.

Most empirical applications of discrete games fall into one of two frameworks. Under the complete information framework, each agent's payoff function is perfectly known to his opponents but not the econometrician. In the incomplete information framework, each agent receives a private shock that is unknown to both opponents and the econometrician. However, in many markets, firms have some private information about their own payoffs—e.g., their contracts with intermediate goods suppliers are private—while at the same time, some determinants of firm demand are commonly observed by the game's players but not by the econometrician—e.g., brand quality of

each firm, as in Berry et al. (1995). To see one implication of abstracting away from either public or private unobservables, consider two firms contemplating entry into a market where only one can operate profitably. If the firms are uncertain about each other's actions, the market may go unserved even though it is profitable for a single entrant. In contrast, a pure strategy complete information model will predict that a firm will always enter the market whenever a monopolist can make positive profits and will interpret neither firm entering as a sign that the market is unprofitable even for a monopolist.

Of course, a key question is whether the data can be used to distinguish between information structures. I will show that the degree of complete versus incomplete information is set identified. For some intuition, assume the econometrician and all players observe a variable, such as a firm-specific cost shifter, that directly affects the profits of only one firm in the market. Differences in rival firms' response to variation in this variable can be used to learn about the degree of complete versus incomplete information. In a complete information game with pure strategies, firms know their opponents' actions when making their own entry decisions. In this case, variation in the cost shifter should not directly affect rival firms' actions because their information set includes opponents' true action. In contrast, when firms have incomplete information, rivals base their decisions on *expectations* of opponents' actions conditional on their information set rather than their true action. In this case, variation in the cost shifter will directly impact rivals' actions because it shifts expectations about the entry of the initial firm. However, because the firms actual action is not in rivals' information sets, it cannot directly affect rivals' actions. Hence, the relative degree of correlation between the rival's action, the firm's action, and the firm's cost shifter contains information about the relationship between public and private information that can be exploited by a structural model.

The flexible information structure I propose includes both publicly observed and privately known structural errors for each firm. A complicating feature of this model, common to many discrete game models, is that it admits multiple equilibria. I allow for a non-parametric equilibrium selection mechanism which imposes no additional restrictions on outcomes beyond equilibrium. In particular, the model allows for the possibility that multiple equilibria are played within the observed data.

I examine identification of the model parameters under this framework. As in other models of discrete-choice games, the parameters of player payoff functions are point identified if the model covariates have sufficiently rich support. Although the information structure parameters are not point identified, the extreme assumptions of pure complete and incomplete information are testable. When the rich support assumption does not hold, as occurs in my application, all the model parameters are set identified. I derive the identified set of the parameters based on the model likelihood and show how to conduct inference on the model parameters whether or not the model is point identified.

The empirical results illustrate that strong assumptions on the information structure have real consequences for the analysis of games. For example, I find that the incomplete information model is excluded from the confidence region implied by the full model, while a pure complete information parameterization is inside the 95 percent confidence region of the full model. However, the confidence region also includes parameterizations where over half the variance is generated by the private information shock. Merely testing the extremes of pure complete and pure incomplete information would ignore this intermediate outcome. Moreover, although the pure strategy complete information framework is not rejected, counterfactual analysis shows that bounds produced using this framework are driven by assumptions on the information structure. In short, without knowing the results of the flexible model, researchers who impose strong informational assumptions would draw overly strong conclusions about policy relevant statistics.

By allowing for both a public and private structural error, I unify two strands of the literature on discrete-choice games. Papers employing complete information games include Bjorn and Vuong (1984); Bresnahan and Reiss (1990, 1991a,b); Berry (1992); Mazzeo (2002); Davis (2006); Tamer (2003); Bajari et al. (2010); Beresteanu et al. (2008), and Ciliberto and Tamer (2009). Many of these papers emphasize the existence of multiple equilibria and propose various approaches to confront the problem. Recently, models of incomplete information, first proposed for empirical work by Seim (2006), have been popular in part because of their relative tractability. Applications and extensions of the incomplete information framework include Augereau et al. (2006); Sweeting (2009); Aradillas-Lopez (2010); Aguirregabiria and Mira (2007); Bajari et al. (2007, 2010b); Paula

and Tang (2012) and Vitorino (2012).

This article conducts sharp inference in a game with public and private unobservables without strong assumptions on equilibrium selection. Tamer (2003) showed that complete information games were identified under rich support assumptions. More recently, Galichon and Henry (2011) have shown how to conduct sharp inference in complete information games with pure or mixed strategies, but do not consider private information. In independent work, Beresteanu et al. (2011) propose a general method of sharp inference whenever the predictions of the model can be described as a convex set of moment conditions. In an appendix, they show the model of this article fits their framework. In contrast to Beresteanu et al. (2011), I use the model likelihood to characterize the identified set, then apply the sieve profile likelihood method proposed by Chen et al. (2011) to conduct sharp inference.

The particular application also contributes to the growing literature on the effect of supercenters on traditional grocery stores. Several studies have investigated how traditional grocery stores are affected by supercenter entry (e.g., Singh et al., 2006; Hausman and Leibtag, 2007; Basker and Noel, 2009; Chiou, 2009; Haltiwanger et al., 2010; Matsa, 2011; Beresteanu et al., 2010; Ellickson and Grieco, 2013). Two things are distinctive about my empirical setting. First, I examine the binary decision of whether or not to operate a firm within a given market, and abstract away from decisions on how to operate that firm, such as determination of prices, product variety, product quality, or location within a market.¹ As consumers face a relatively small choice set when choosing where to do their grocery shopping, the effect of firm openings and closings may have a much larger impact on consumer welfare than competitive responses in product price or quality. Second, I focus on rural grocery markets only. Due to the small number of stores in rural markets, these markets are most impacted by closure of a grocery store. These choices serve to focus the analysis on the commonly voiced complaint that supercenters harm consumers because they crowd out traditional grocery stores.

Supercenters are commonly believed to hold a significant cost advantage over traditional grocery

¹In a related paper Beresteanu et al. (2010) examine the opening and closing decisions of chain grocery stores within grocery distribution markets. While they focus on large chain stores who compete over a wide area, this article abstracts away from chain effects and focuses on small stores—many of which are single-store firms—in more isolated rural areas.

stores due to their scale and integrated distribution networks. However, the locational convenience of a local grocery store may provide some insulation from competition with supercenters. The magnitude of insulation is the key empirical question of this article. To answer this question, I use data on entry and exit in rural grocery markets from 1998 to 2002 to estimate an entry game model that allows for the calculation of expected long-run firm values.

Rural grocery markets have features of both complete information and incomplete information models. There are differences across markets in local terrain, zoning regulations, local tastes, and location availability, as well as differences among firms in the quality of their products and the level of customer service which are observable to all players but not the econometrician. However, firms and potential entrants also have some cost information that is kept private from their competitors and the econometrician. For example, firms have private information regarding their management expertise, outside opportunities, and the ease of integrating into a distribution network. By allowing both a public and private error term for each player, my approach is able to account for both of these features.

The empirical results illustrate that the role of supercenters in local grocery markets is less dramatic than is commonly thought. Entry by a supercenter outside, but within 20 miles, of a local monopolist's market has a smaller impact on firm profits than entry by a local grocer. Although supercenters appear to be associated with a decrease in stores' expected profits, and appear to lower the number of grocery stores in surrounding markets, the effects are small. I interpret this as evidence that location and format-based differentiation partially insulates rural stores from competition with supercenters. Indeed, when accounting for supercenters' impact on opponent behavior, I cannot reject the possibility that in equilibrium, supercenters *increase* the value of local grocery monopolists through discouraging local entry. In conclusion, I find a small reaction in the number of expected local stores to supercenter entry, particularly in monopoly markets. Although I do not directly model consumer demand, the small crowding-out effect documented by the model implies that welfare losses due to fewer small stores is unlikely to offset the welfare benefits of supercenters through increased price competition and variety that are highlighted by Hausman and Leibtag (2007) and other authors.

The article is organized as follows. Section 2 introduces the model. Section 3 provides a discussion and uses numerical examples to illustrate how the model incorporates public and private information. In Section 4, I study the identification of the model. Section 5 shows how to conduct inference. Section 6 introduces the empirical application and the data. The results of the application of the full structural model and several counterfactual experiments are presented in Section 7. The final section concludes by reiterating that allowing for flexible information structures improves the credibility of empirical investigations of discrete games. Proofs are presented in Appendix A.

2 The Model

Consider a small market with two potential entrants $i \in \{1, 2\}$.² The firms simultaneously choose whether or not to be active, where $y_i = 1$ if firm i is active and $y_i = 0$ otherwise, so the observed action profile is $y = (y_1, y_2)$. In order to concentrate on the role of the information structure, I abstract away from dynamics and assume firms make a one-shot long-run entry decision. Long run payoffs are described by the following linear reduced form payoff function,

$$\pi_i(y_i, y_{-i}; x, \theta) = \begin{cases} x_i\theta_{i\mu} + y_{-i}x_i\theta_{i\delta} + \epsilon_i + \nu_i & \text{if } y_i = 1 \\ 0 & \text{if } y_i = 0 \end{cases}. \quad (1)$$

The researcher observes $x = (x_1, x_2)$ which may contain common components (e.g., population) or firm-level components (e.g., firm age or earlier capital investment). The term $x_i\theta_{i\mu}$ captures the impact of the firm and market characteristics on monopoly profits, whereas the term $y_{-i}x_i\theta_{i\delta}$ captures the competition effect of a second firm operating on the profits of firm i . The competition effect may depend on observable firm and market characteristics. This specification is in line with Bresnahan and Reiss (1990) who provide a simple model of entry to show that the impact of entry may be decomposed into a fixed component and a component which varies with market size. The parameters of the payoff functions are $\theta_p = (\theta_{1\mu}, \theta_{1\delta}, \theta_{2\mu}, \theta_{2\delta})$.³

²Extending the model to n players is conceptually straightforward but computationally burdensome due to the need to compute all Bayesian Nash equilibria of the game to compute the likelihood function. Because my application will involve rural markets with two players, I follow Bresnahan and Reiss (1991a) and Tamer (2003) in focusing on two-player games.

³A simple extension would be to allow the competition effect to depend on observables or unobservables. The

Each players' payoff includes two components which are unobserved by the econometrician. The public shocks, $\epsilon = (\epsilon_1, \epsilon_2)$, are jointly observed by both players in the game. I assume they are bivariate normal with a common variance component σ_ϵ^2 and a correlation coefficient ρ . The public shock captures the drivers of firm profit which are unobserved by the econometrician but known to all players, (e.g., firm location, local tastes, etc.). In addition to the public shock, each firm receives a private shock ν , to capture the impact of drivers of profit which are known only to firm i (such as management capacity or opportunity costs) drawn from a normal distribution with variance σ_ν^2 . Independence of ν_i conditional on x and ϵ is all that is necessary to ensure that player i 's beliefs about equilibrium play are not dependent on ν_i and that optimal strategies are monotone cutoffs in ν_i , however I assume unconditional independence for tractability within my applied setting.⁴ Because only discrete outcomes are observed, a scale normalization is necessary to identify the model. I normalize the total variance of the combined structural error, $\sigma_\epsilon^2 + \sigma_\nu^2$ to be one. Therefore, the parameters of the information structure are $\theta_e = (\rho, \sigma_\epsilon^2)$. The full set of parameters to estimate are $\theta = (\theta_p, \theta_e)$.

Equilibrium

Players choose actions using equilibrium strategies. From the perspective of the players, who observe ϵ , the model is a game of incomplete information as long as $\sigma_\nu > 0$.⁵ I confine the analysis to pure strategy Bayesian Nash equilibria—a mapping from the firms' private types ν to an action profile $y \in \{0, 1\}^2$.⁶

partial identification results presented in Section 4 and the inference technique presented in Section 5 trivially extend to this approach. However, the point identification results in Section 4 require sign restrictions on the competition effect which would be cumbersome with a more general competition effect.

⁴Most studies of incomplete information games assume that agents' private types are independent, though there are three exceptions. Aradillas-Lopez (2010) allows for correlation when beliefs are of the form $P(y_{-i} = 1|x, y_i)$ rather than $P(y_{-i} = 1|x, \nu_i)$. Wan and Xu (2012) and Liu et al. (2012) develop identification results assuming players' strategies are monotone in ν_i .

⁵If $\sigma_\nu = 0$, then the game is of complete information, however the famous result of Harsanyi (1973) has shown that firm's mixed-strategy equilibrium actions are the limit of the sequence of pure-strategy Bayesian Nash equilibria of incomplete information games as $\sigma_\nu \rightarrow 0$. For expositional purposes, I concentrate on the $\sigma_\nu > 0$ case knowing that I must take limits for $\sigma_\nu = 0$.

⁶Milgrom and Weber (1985) have shown that in incomplete information games such as the one considered here where player types are conditionally independent, equilibrium must exist and every mixed strategy equilibrium has a nearby "purification" pure strategy under which the agents distribution of observed behavior and expected payoffs are identical.

A firm's optimal strategy is a cutoff in ν_i . Increasing ν_i unambiguously increases the expected profits of entry (action 1), so a strategy in which the firm operates at $\nu_i' < \nu_i''$ but not at ν_i'' is clearly sub-optimal. Therefore, an optimal strategy must be of the form,

$$s_i(\nu_i, x, \epsilon; \theta) = \begin{cases} 1 & \text{if } \nu_i \geq \chi_i(\epsilon, x; \theta) \\ 0 & \text{otherwise} \end{cases},$$

where $\chi_i(x, \epsilon; \theta)$ is the entry cutoff for agent i . This is convenient because optimal strategies are associated with their cutoffs, which are real numbers rather than functions. Players know the distribution of ν_i , so their beliefs about their opponents probability of entry given their opponent's strategy are,

$$\varrho_i(\chi_i, \epsilon, x; \theta) = \int s_i(\nu_i, \epsilon, x; \theta) d\Phi(\nu_i) = 1 - \Phi\left(\frac{\chi_i(\epsilon, x; \theta)}{\sigma_\nu}\right) \quad (2)$$

Players select their cutoffs to optimize their expected payoff given their beliefs about the actions of other players. Let $\chi_i^b(\chi_j, \epsilon; \theta)$ denote player i 's best response to player j playing the cutoff χ_j . Player i 's best response is to adopt the cutoff where he is exactly indifferent between his two actions. i.e, firm i 's best response is to operate when $\nu_i \geq \chi_i^b(\chi_j, \epsilon; \theta)$. The following equation defines agent i 's best-response cutoff, $\chi_i^b(\cdot)$, as a function of opponent strategies and the publicly observed ϵ :

$$\chi_i^b(\chi_j, \epsilon, x; \theta) = -(x_i\theta_{i\mu} + \varrho_j(\chi_j, \epsilon; \theta)x_i\theta_{i\delta} + \epsilon_i) \quad (3)$$

Any joint set of cutoffs $\chi = (\chi_1, \chi_2)$ that satisfy (3) for both players represents an equilibrium.⁷ Because there is a simple one-to-one mapping between χ_i and ϱ_i , we can alternatively describe the equilibrium in terms of either cutoffs or entry probabilities.

⁷If there were N players in the game, (3) would need to be modified to account for each opponents probability of entry, this would result in a system of N equations and N unknowns which would jointly determine the Bayesian Nash equilibrium of the game. In general, this problem is very computationally burdensome, although see Bajari et al. (2010a) for a novel approach to computing all equilibria in large incomplete information games.

Multiple Equilibria and Equilibrium Selection

Using the necessary and sufficient conditions derived in the previous section, I can characterize the equilibrium set for the incomplete information game as the solution set to a system of nonlinear equations.

$$\mathcal{E}(\epsilon, x, \theta) = \{\chi : \forall i, \chi_i = \chi_i^b(\chi_{-i}, \epsilon, x; \theta)\}$$

Each equilibrium implies a multinomial distribution across action profiles. If the equilibrium set were a singleton everywhere, a unique observable outcome distribution for the model could be obtained by integrating over the public information shock ϵ . However, when there are multiple equilibria, the model does not provide a unique distribution over actions, so the model is incomplete (Tamer, 2003).

If more than one equilibrium exists, one is selected from $\mathcal{E}(\epsilon, x, \theta)$ on the basis of a public coordination device that may depend on (ϵ, x, θ) but is independent of ν . This is a consequence of the assumption that an equilibrium profile is played, not an additional assumption. The public coordinating device must be independent of ν (like all other variables in the model) to prevent information leakage, which would cause agents' beliefs about other players to depend on private information. Once players observe the device, it is clearly not optimal for any player to unilaterally deviate to play any strategy other than the one selected.

A selection mechanism is a function that maps the space of strategy profiles to the probability that a particular strategy profile is played. To be valid, a selection mechanism only plays strategy profiles which are equilibria with positive probability.⁸

Definition 1. *Let $\lambda : \mathbb{R}^2 \times X \times \Theta \rightarrow [0, 1]^{\bar{E}}$ be a selection mechanism, where $\lambda^e(\epsilon, x, \theta)$ is the probability that strategy profile e is played when the game is defined by (ϵ, x, θ) and \bar{E} represents the maximum number of equilibria in the model.⁹ Let Λ be the set of valid selection mechanisms, where a selection mechanism is valid if:*

⁸Conditional on the publicly observed information, strategies can be described as a vector of cutoffs, such that in a two-player game, the selection mechanism is a function from \mathbb{R}^2 to $[0, 1]$.

⁹This number is generically finite. Moreover, when all errors are normal and there are two players $\bar{E} = 3$. This result can be shown using the fact that best-response functions for this game are translations of the normal distribution function.

1. If $e \notin \mathcal{E}(\epsilon, x, \theta)$, $\lambda^e(\epsilon, x, \theta) = 0$,

2. $\sum_{e \in \mathcal{E}(\epsilon, x, \theta)} \lambda^e(\epsilon, x, \theta) = 1$.

In order to enforce validity constraints, the selection mechanism is a function of the model parameters, θ . The equilibrium selection mechanism is an infinite dimensional parameter that completes the model. In other words, given (θ, λ) , the model predicts a unique probability distribution over actions and the model likelihood is well defined. If only θ is specified, then a set of probability distributions are available and the model is only partially determined. The inference procedure proposed in Section 5 will maximize over this set to determine whether the model parameters θ are consistent with the data.

3 Impact of Information Structure

The model combines the complete and incomplete information frameworks and nests many models presented in the literature. After a scale normalization, the information structure is parameterized by two variables: the relative variance of the publicly observed structural error, σ_ϵ^2 and the correlation between firms' public errors, ρ . The earliest entry games assumed that markets were subject to a single profit shock that is homogeneous across firms and publicly known to the firms (Bresnahan and Reiss, 1991b). The flexible model of this article nests this model when $\sigma_\epsilon^2 = 1$ and $\rho = 1$. The more general complete information model, which I take as the baseline complete information framework, allows for firm-level heterogeneity in the error term. This is equivalent to assuming $\sigma_\epsilon^2 = 1$ within my full model, but allowing ρ to vary.¹⁰ The pure incomplete information framework is attained within the full model when $\sigma_\epsilon^2 = 0$, eliminating the public structural error from the model. A market-level public shock could be added to an incomplete information game, which would be accomplished within my framework by assuming that $\rho = 1$ and allowing σ_ϵ to vary.¹¹ Before

¹⁰Most applications of the complete information model have restricted themselves to pure strategy equilibria (e.g., Berry, 1992; Mazzeo, 2002; Ciliberto and Tamer, 2009), so I also take the pure strategies assumption as part of the complete information framework.

¹¹Seim (2006) uses a two-stage game where potential entrants first decide whether to enter based on a publicly observable market-level shock, and then decide on their location given the number of firms that entered. The location decision is an incomplete information game. The two-stage structure is distinct from simply assuming $\rho = 1$, as in the model presented here.

turning to identification, it is important to understand how differences in the information structure will impact observable outcomes. Section 3 examines how restricting the information structure can affect empirical results. Section 3 illustrates how the information structure affects the measure of markets that have multiple equilibria and the shape of the equilibrium correspondence.

The different effects of uncertainty and heterogeneity

The two structural shocks, ϵ and ν , affect strategies and observed action distributions in different ways. The key difference is that ϵ_2 affects the expectations of player 1 whereas ν_2 does not. Abstracting away from either shock simplifies the model in ways that will impact estimation and policy analysis.

First consider the case where $\sigma_\epsilon = 0$, so only private information is present. Under the assumptions of the incomplete information framework, econometricians often assume they can directly measure equilibrium strategies in the data by inverting conditional choice probabilities (e.g., Bajari et al., 2010b).¹² By assumption, all firms facing the same unobservables have the same expectations about their opponents' actions. If there is even a small amount of public information unobserved by the econometrician, this estimate of strategies will bias the degree of player uncertainty upwards because the private error is the only source of unobserved heterogeneity in the model. This necessarily biases firms' expected profits, often an object of counterfactual interest, downwards.

Abstracting away from private information also imposes severe restrictions on the model's implications. Under the complete information framework, firms only have uncertainty about their opponents if mixed strategy equilibria are being played. Most empirical work using complete information has also assumed pure strategy equilibria are played, meaning that firms are able to completely avoid coordination failure and the resulting negative profit outcomes. As a result, the pure strategy complete information model ignores the potential impact of coordination failure in policy counterfactuals, even though, when incomplete information is present, changes in parameters may in practice result in an increase or decrease in the probability of such a coordination failure.

¹²This requires the additional assumption that all markets with the same covariate choose to play the same equilibrium.

Relationship between Information Structure and Multiple Equilibria

The extent to which multiplicity causes an identification problem is related to the *proportion* of markets with multiple equilibria. How does the amount of public versus private information affect the propensity for multiple equilibria? To investigate this question I present a numerical example calculating the region of multiplicity as the variance of the private error changes. For this exercise, I consider a simplified symmetric version of the model with no covariates, so θ_μ and θ_δ are now scalars, the profit function for firm i is simply,

$$\pi_i(y_i, y_{-i}; x, \theta) = \begin{cases} \theta_\mu + y_{-i}\theta_\delta + \epsilon_i + \nu_i & \text{if } y_i = 1 \\ 0 & \text{if } y_i = 0 \end{cases}$$

Figure 1 presents a graph indicating the region of multiplicity in ϵ -space for the model where $\theta_\mu = 0.5$ and $\theta_\delta = -1$. This figure can be thought of as a generalization of Figure 1 in Bresnahan and Reiss (1991a) who present the region of multiplicity within the complete information framework. The variance of the incomplete information shock, σ_ν , is increasing across the four panels of the figure. The distribution of ϵ establishes determine the density of markets across ϵ -space, but plays no direct role in Figure 1 itself. Note that determining the proportion of markets with multiple equilibria involves integrating over the space illustrated in Figure 1.¹³ The limit result of Harsanyi (1973) is apparent: as σ_ν becomes small, the region of multiplicity closely resembles the “box” of multiplicity in the complete information game studied by Bresnahan and Reiss (1991a) and Tamer (2003).

Figure 1 illustrates one particular parameterization of the model. Two striking observations can be drawn from it. First, the size of the region of multiplicity tends to decrease as σ_ν increases, and eventually vanishes. Below I discuss why an increase in the degree of uncertainty is likely to limit the scope of multiple equilibria.¹⁴ Second, multiplicity in the presence of uncertainty is much

¹³Figures 1 and 2 describe the changes in the equilibrium set from the prospective of players (who observe ϵ and the payoff parameters). In the empirical model, the total variance of $\epsilon_i + \nu_i$ is normalized to one (the econometrician must integrate out ϵ); in these figures ϵ and the profit function parameters are taken as given. With this information, the model provides distinct predictions for $\sigma_\nu \in [0, \infty)$.

¹⁴In a different, but similar, context, Morris and Shin (2000) have argued that adding incomplete information can reduce the degree of multiplicity in a model. However, in general, adding uncertainty can increase or decrease the number of equilibria depending on the game. Figure 2 provides an example where the number of equilibria rises from

more likely when the two firms are similar in terms of their publicly observed propensity to enter (e.g., markets near the line $\epsilon_1 = \epsilon_2$ continue to exhibit multiplicity even when σ_ν is relatively high). This result relates to the strategic substitutability of player actions when δ is negative. When the game is of strategic complements, i.e., $\theta_\delta > 0$ the relationship is reversed, and multiplicity is most prevalent along the $\epsilon_1 = -\epsilon_2$ line.¹⁵ Although these figures do not constitute a proof, the results are stable across several different parameterizations of the model.

Intuition for the decrease in the size of the region of multiplicity can be found by considering players' probability of entry given ϵ and x . Figure 2 displays how equilibrium entry probabilities vary with σ_ν for four different public error outcomes. In subplots (a) and (b), firm 1 is profitable regardless of firm 2's action when $\sigma_\nu = 0$. For plots (c) and (d) there are multiple equilibria at $\sigma_\nu = 0$ as firm 1's profitability depends on firm 2's action.

When σ_ν increases from zero, entry probabilities can no longer be degenerate (although they remain close to degenerate). This decrease causes firm 2 to adjust its expectations of firm 1 entry, thus adjusting its own entry cutoff. Multiple equilibria may be sustained because 3 different sets of cutoffs jointly solve the equilibrium conditions. The extent to which expectations affect firms' optimal cutoffs is limited (expected entry probabilities are bounded between 0 and 1). As σ_ν becomes large, the probability that a firm's optimal action is independent of its expectations (e.g., ν_i is such that the optimal action is to enter even if the opposing firm will enter with certainty) increases. Hence, the set of rational strategies shrinks, forcing equilibrium strategies to be more similar. In the limit as σ_ν approaches infinity,¹⁶ whether or not the firm expects entry is irrelevant to its entry probability. Given each firm has only one rationalizable entry probability there is of course a unique equilibrium

The result that private information can dramatically reduce the amount of multiplicity within the model suggests that—even without specifying the mixing distribution across equilibrium—the

1 to 3 and then falls back to 1 as the degree of uncertainty increases.

¹⁵Of course, when $\delta = 0$, there is no strategic interaction and there is always a unique equilibrium.

¹⁶In the empirical model, the econometrician must integrate over ϵ and normalize the variance of $\epsilon_i + \nu_i$. In Figures 1 and 2, we have fixed ϵ and the payoff parameters, so σ_ν can vary from 0 to ∞ (see Footnote 13). A change in σ_ν , is isomorphic to holding the variance of σ_ν fixed, and multiplying the payoff parameters and ϵ by a constant $c \in [0, 1)$. This should make the limit argument clear: Taking $\sigma_\nu \rightarrow 1$ is analogous to taking $c \rightarrow 0$. When payoff parameters and public shocks are all zero, it is trivial to see the game has a unique equilibrium.

bounds on entry probabilities under a parameterization that admits a moderate amount of private information may be much tighter than the bounds of a similar complete information parameterization. This is welcome news, as the generalization from complete information to the full model is unlikely to exacerbate the identification issue related to multiplicity. However, because the model is at the height of its flexibility under the complete information assumption—multiple equilibria is most common and different equilibria allow the widest range of equilibrium entry probabilities—it may be difficult to reject the complete information framework.

4 Identification

The full model nests the complete and incomplete information frameworks as endpoints of a continuum of possible information structures. I study identification under rich support assumptions in Section 4. Tamer (2003) shows that the payoff function of a complete information game is point identified if the covariates have a rich support, and I show that this result extends to the general model. Although I am unable to establish point identification of the information structure itself, I show that both the complete and incomplete information assumptions commonly used in the literature are testable against the general framework.

Without assuming that some covariates have rich support, even the payoff parameters of the model are only partially identified. In Section 4, I derive the identified set for the model parameters without assuming a rich support.¹⁷ Even though the model is set identified, this may have little practical effect on the results of the estimation if the identified set is small.¹⁸ Section 5 will provide techniques to perform inference on the model and conduct counterfactual analysis using confidence regions for the parameters of interest without assuming point identification.

¹⁷Bajari et al. (2011) study a similar model and argue that it may not be point-identified without parametric restrictions on the selection mechanism. In Section 4, I explore what can be learned from the model without imposing point identifying assumptions.

¹⁸See Honoré and Tamer (2006) for an illustration of this point in the context of a single-agent dynamic model.

Identification with Rich Support Assumptions

In this section, I analyze the model under the assumption that the covariates have a rich support. With this assumption, I first show that the parameters of firm payoff functions in the general model are point identified. Intuitively, this is because in the limit, the econometrician can observe markets where a firm's probability of entry given its covariates is 1 or 0 regardless of the realization of the publicly observed shock. In these cases, agents always take opponent entry as given, and the problem reduces to a standard threshold crossing model. I then show that the implications of both the complete information model and the incomplete information model are testable.

First, I state the basic assumptions.

Assumption 1. *An econometrician observes a random sample of M markets $\{(y_m, x_m)\}$, $m = (1, \dots, M)$, where $y_m = (y_{1m}, y_{2m})$ are binary indicators of whether firm i operates in the market and $x_m = (x_{1m}, x_{2m})$ are vectors of covariates.¹⁹*

Assumption 1 merely guarantees that we can consistently estimate the conditional probabilities $P(y|x)$.

Assumption 2. *For $i \in \{1, 2\}$, there is of at least one covariate, x_i^* that is not included in x_{-i} , and has positive density everywhere on \mathbb{R} conditional on all other covariates. The coefficients of the payoff function associated with x_i^* satisfy $|\theta_{i\mu}^*| > |\theta_{i\delta}^*|$.*

Rich support assumptions are a common tool in the identification of games. The assumption used here is similar to that of Tamer (2003), who analyzes identification in the context of a pure strategies complete information game. Theorem 1 generalizes the result of Tamer (2003) for complete information games to the general model with an unspecified information structure.

Theorem 1. *If Assumptions 1 and 2 hold, the parameters of the payoff function $(\theta_{i\mu}, \theta_{i\delta})$ for $i = 1, 2$ are point identified.*

The argument of the proof shows that the payoff parameters can be identified by observing those markets where strategic interaction is not a factor. In particular, consider markets where

¹⁹I suppress the market subscript when it is clear from the context.

x_i^* is very low or very high. In such a markets, firm i 's decision is essentially known: regardless of the outcome of ϵ_i or ν_i , she will enter with probability 1 or 0, depending on the effect of x_i^* on profits. In this case, both player $-i$ and the econometrician know what firm i will do, and the econometrician can use variation in x_{-i} to estimate the payoff parameters of firm $-i$.²⁰ The procedure can be reversed to estimate the payoff parameters of firm i .

Notice that, in the markets which provide identification, one firm has a dominant strategy to enter or not enter. Therefore, neither multiplicity of equilibria nor the information structure are an obstacle to identification of the payoff parameters. However, for the same reason, there is no way to use these markets to identify anything about either parameter of the information structure. Intuitively, the rich support assumption aids identification of the payoff parameters because it ensures that there are markets where one firm has a known dominant strategy (always enter or never enter). However, precisely because one firm's action is constant in these markets, they provide no information about either the correlation between firm's public errors, ρ , or the relative amount of public versus private information, σ_ϵ^2 . Alternatively, for markets where neither firm has a dominant strategy, it is impossible to rule out multiplicity of equilibria as a determining factor for the observed correlation between firms' actions. In this case, actions are a function of both the information structure and the equilibrium selection mechanism.

Although the information structure cannot be point identified in general, it is possible to test some commonly used restrictions on the information structure. The next theorem shows that assumptions of the commonly used pure strategy complete information model can be tested against the general model if the following assumption holds. To do so, the following assumption ensures that there exists a pure strategy equilibrium under complete information.

Assumption 3. *The competition effects $x_i\theta_{1\delta}$ and $x_i\theta_{2\delta}$ have the same sign for all x .*

This assumption is necessary only when testing the complete information framework as it does not make sense to test the pure strategies complete information model when it is incoherent and does not admit a solution. An implication of this assumption is that $\theta_{i\delta}^* = 0$ for the rich support

²⁰The competition effect itself is identified by comparing the case where firm i 's dominant strategy is to always enter (e.g., $x_i^* \rightarrow \infty$) with the case where her dominant strategy is to never enter ($x_i^* \rightarrow -\infty$).

parameters, which is frequently assumed in empirical work. The full model allows for mixed strategies under complete information, so this assumption is needed only for Theorem 2. Because $\theta_{1\delta}$ and $\theta_{2\delta}$ are point identified, Assumption 3 is easy to test.

Theorem 2. *If Assumptions 1 through 3 hold, the pure strategies complete information framework implies testable restrictions on $P(y|x)$.*

The intuition of the proof is to find a set of restrictions which must hold under complete information. Complete information models assume that all unobserved variation is observable to all players, eliminating ν from the model while retaining ϵ . As Theorem 1 showed, the payoff parameters are identified. Under the assumption of complete information the only parameters remaining to be identified are the correlation of the public shock, ρ , and the selection mechanism $\lambda(\cdot)$. The proof of Theorem 2 provides a set of equality restrictions for a sub-vector of $P(y|x)$ which involve ρ , but not $\lambda(\cdot)$, with which to test the pure strategies complete information assumption. These restrictions rely on the fact that, under complete information, some outcomes only occur when the equilibrium is unique, so $\lambda(\cdot)$ does not affect those outcomes' probabilities. For example, assume that duopoly profits are strictly lower than monopoly profits.²¹ Then, if duopoly is an equilibrium, both firms are playing dominant strategies.²² By examining the probability of this outcome at a particular value of x , there are multiple equations available to identify the single coefficient, ρ . However, there are an infinite number of such restrictions (one for each observed value of x). If they cannot all be jointly satisfied, the complete information assumption on the information structure can be rejected.

The key feature of these equalities is that they do not depend on the selection mechanism, λ , which is not identified. This is also the reason why the test must focus on pure strategies equilibria. If mixed strategies were allowed, a (non-degenerate) mixed strategy from a game with multiple equilibria would result in *any* outcome occurring with positive probability, introducing λ into every outcome probability expression. Conveniently, almost all applications of complete information games assume only pure strategy equilibria are played, so Theorem 2 is relevant to

²¹This is the most natural case, but Assumption 3 is slightly weaker.

²²This theorem could easily be extended to a game with N players, but it would rely on “extreme” outcomes of all entering or all not entering for to construct the test.

empirical work.²³

A similar test can be devised for the incomplete information model, the other commonly employed information structure.

Theorem 3. *If Assumptions 1 and 2 hold, the pure incomplete information framework implies testable restrictions on $P(y|x)$.*

Like the test of complete information, incomplete information framework can be tested by concentrating on equality restrictions for $P(y|x)$ that do not involve the selection mechanism. In the incomplete information case, $\sigma_\epsilon = 0$ and ρ drops out of the model as there are no publicly observed profit shocks. In this case, Theorem 1 identifies all parameters of the model except $\lambda(\cdot)$. Therefore, the econometrician calculate the equilibrium set for every x and determine its cardinality. Where there is a unique equilibrium, this exactly pins down $P(y|x)$. If these equalities do not hold across all x for which equilibrium is unique, the incomplete information framework can be rejected. Furthermore, when there is a unique equilibrium, it must be that players' actions will be independent conditional on the other observables, as all variation is due to the private shocks.²⁴ The conditional independence requirement is very strict and, thus, it would appear that the incomplete information model will be rejected in many cases, as happens in the empirical analysis presented below.

The general model does not appear to be point identified. In essence, both of the above tests focus on the degree of correlation of player actions conditional on x in situations where selection can be ruled out as an explanation for that correlation.²⁵ If $\sigma_\epsilon \in (0, 1)$, there is no set of restrictions that isolates changes in σ_ϵ from changes in λ . This means that the reason for correlation patterns in player actions may be attributed to either the selection mechanism or the degree of correlated

²³The full model does allow for mixed strategies within the complete information context, so the test implied by Theorem 2 is actually stronger (more likely to be rejected) than simply testing $\rho = 0$ in the full model (which considers the complete information framework with or without mixed strategies).

²⁴A recent paper on multiplicity in games of pure incomplete information uses this fact to develop a test for multiplicity based on conditional independence (Paula and Tang, 2012). When multiple equilibria are present, they further show that bounds restrictions are available.

²⁵In the complete information case, we are directly testing the probability of (1,1) and (0,0) outcomes. In the incomplete information test, we are able to test the entire probability distribution when equilibrium is unique, but a key source of power in this test is that player actions are conditionally independent under the incomplete information framework when equilibrium is unique.

public information. In principle, one could use the model selection test developed by Shi (2011) to directly compare the extreme cases of pure complete and incomplete information. However, there is no reason to assume that only one type of heterogeneity is present. Instead, I make use of the fact that the scope of multiplicity to affect the outcome distribution is bounded in two ways: First, selection only plays a role when there are multiple equilibria, and for many parameter values, the degree of multiplicity may be small. Second, even where there is multiplicity, the outcome distribution must be a valid mixture between equilibria, limiting the extent to which the selection mechanism can induce correlation across outcomes. The following section uses these observations to show how to construct sharp bounds for the identified set of the richer model with both complete and incomplete information shocks.

Identified Set

In empirical work, most datasets have only a limited amount of variation in their covariates; the rich support assumption is often too strong. For this reason, I study identification under the following assumption,

Assumption 4. *For $i = \{1, 2\}$, the covariates $x_i \in X$ are discrete.*

The discrete support assumption is the opposite extreme from the rich support assumption of the previous section.²⁶ Under this assumption, I derive the identified set of the model. In Section 5, I use this result to infer a confidence region for the model’s parameters.

For expositional simplicity I first consider in Section 4 identification in the context of an incomplete information game where there is no public information shock. In Section 4, I extend this result to derive the identified set for the general model with both complete and incomplete information structural errors.

Identified Set with Only Incomplete Information

For this section only, assume that there is no public shock, so $\sigma_\epsilon = 0$ and $\nu_i \sim N(0, 1)$ is independent of x and i.i.d. across players. As $\epsilon_i = 0$ under this assumption, I temporarily drop it from

²⁶This section could easily be extended to continuous but bounded regressors.

the notation. Section 2 derived the necessary and sufficient condition equilibrium strategies and entry probabilities for this model (the only difference is that ϵ is eliminated from the equilibrium constraints). An equilibrium can be expressed either as a vector of cutoffs $(\chi_1(x, \theta), \chi_2(x, \theta))$ or entry probabilities $(\varrho_1(x, \theta), \varrho_2(x, \theta))$, which are related through the one-to-one mapping (2). It is more convenient to describe the equilibrium in terms of the entry probability profile $\varrho(x, \theta)$. As shown above, this model may have multiple equilibria; hence, the solution to the model is a correspondence $\mathcal{E}(x, \theta)$ that maps a set of covariates and the parameters to a finite set of equilibrium strategy profiles. Let $\#\mathcal{E}(x, \theta)$ be the cardinality of this set and let e index an element of $\mathcal{E}(x, \theta)$ such that $\varrho^e(x, \theta)$ is a vector of choice probabilities related to equilibrium e .

Every equilibrium profile implies a multinomial distribution over outcomes. For a profile $e \in \mathcal{E}(x, \theta)$, let $\tilde{\Psi}(y|x, \theta, e)$ be the resulting outcome probabilities.²⁷ The probability of observing outcome y if e is the selected equilibrium is,²⁸

$$\tilde{\Psi}(y|x, \theta, e) = \prod_{i=1}^2 \varrho_i^e(x, \theta)^{\mathbf{1}[y_i=1]} (1 - \varrho_i^e(x, \theta))^{\mathbf{1}[y_i=0]}. \quad (4)$$

If equilibrium were unique, $\tilde{\Psi}(y|x, \theta, e)$ could be compared directly to the observed data $P(y|x)$. When equilibrium is not unique, our assumptions imply that the observed outcome distribution is some mixture of equilibrium strategies according to a valid equilibrium selection mechanism $\lambda(x, \theta)$. For a given selection mechanism, the probability of outcome y is:

$$\Psi(y|x, \theta, \lambda) = \sum_{e \in \mathcal{E}(x, \theta)} \lambda^e(x, \theta) \tilde{\Psi}(y|x, \theta, e). \quad (5)$$

If λ were a parametric function and the model were identified, (θ, λ) could be jointly estimated using the likelihood function implied by (5). Such a strategy is pursued in the complete information context by Bajari et al. (2010). Another common assumption in the context of incomplete information games is that the same equilibrium is always played in observationally equivalent markets. However, economic theory tells us nothing about the selection of an equilibrium to play, so I do

²⁷In this slight abuse of notation e represents the selection mechanism in which equilibrium e is always played.

²⁸The independence of the agents' decisions conditional on x is implied by the fact that the only structural error within the incomplete information model is independent across players and privately observed.

not wish to impose strong restrictions on the selection mechanism. Instead, I allow λ to be any valid mixture across equilibria and derive the sharp identified set implied by the model. That is, I use the model of the equilibrium correspondence given θ and the fact that I know that $\lambda^e(x, \theta) \geq 0$ and $\sum_{e \in \mathcal{E}} \lambda^e(x, \theta) = 1$ to derive the identified set.

Theorem 4. *Given Assumptions 1 (random sample) and 4 (discrete covariates), the sharp identified set of θ for the incomplete information model ($\sigma_\epsilon = 0$) is,*

$$\Theta_I = \left\{ \theta \in \Theta : \forall y, x, \exists \tilde{\lambda} \in [0, 1]^{\bar{E}} \text{ s.t. } \begin{array}{l} P(y|x) = \Psi(y|x, \theta, \tilde{\lambda}), \\ \sum_{e \in \mathcal{E}(x; \theta)} \tilde{\lambda}^e = 1 \end{array} \right\} \quad (6)$$

Where \bar{E} is a constant which represents the largest possible number of equilibria the model admits for any $x \in X$ and any $\theta \in \Theta$.²⁹

To gain some intuition into how (6) restricts the identified set, consider testing whether θ is the true parameter. For a given θ , I can solve the model for $\mathcal{E}(x, \theta)$ for all x . Suppose, for a given x' , $\mathcal{E}(x', \theta)$ is a singleton; the equilibrium mixing distribution is then degenerate and x' provides three equality restrictions with which to test the parameter θ (the fourth outcome probability is redundant due to the adding up constraint). If there are three equilibria in $\mathcal{E}(x', \theta)$, then I can construct a matrix of outcome probabilities, $\Psi(y|x', \theta, e)$ where the rows represent different outcomes y and the columns represent different equilibria $e \in \mathcal{E}(x', \theta)$. If this matrix has full rank, this system of linear equations (5) restricts the set of admissible values for $\lambda(x', \theta)$ to a singleton. In other words, at most one $\lambda(x', \theta)$ can satisfy (5) for all y , and it is only admissible if it represents a valid selection mechanism (e.g., all its elements are positive). I can conduct such a test for all x , if I observe *any* x such that a valid $\lambda(x, \theta)$ cannot be found, then θ must be excluded from the identified set. Clearly, variation in x improves identification as the test of θ must be satisfied for all x for θ to be in the identified set.

²⁹The identified set could be written making use of the definition of an equilibrium selection mechanism defined in Section 2, however I use simpler notation in (6) because it does not include the public information shock, ϵ . From the index theorem it can be shown that the number of equilibria is generically finite. Under the assumptions of my application, $\bar{E} = 3$.

Identified Set of the Full Model

I now return to the full model presented in Section 2 by allowing $\sigma_\epsilon > 0$ and re-introducing ϵ to the notation. As a result, player strategies and equilibrium selection must both be modeled as functions of the publicly observed error ϵ . Therefore, the selection mechanism is now potentially a function of ϵ and, as a consequence, is an infinite dimensional parameter.

The observed distribution of outcomes is now,

$$\Psi(y|x, \theta, \lambda(\cdot)) = \int \sum_{e \in \mathcal{E}(\epsilon, x, \theta)} \lambda^e(\epsilon, x, \theta) \tilde{\Psi}(y|\epsilon, x, \theta, e) dF(\epsilon; \theta). \quad (7)$$

The identified set can now be expressed as follows,

Theorem 5. *Given assumptions 1 and 4, the sharp identified set of θ for the full model is,*

$$\Theta_I = \{\theta \in \Theta : \forall x \in X, \exists \lambda \in \Lambda \text{ s.t. } P(y|x) = \Psi(y|x, \theta, \lambda)\}, \quad (8)$$

where,

$$\Lambda = \left\{ \lambda : \forall \epsilon, x, \theta, \sum_{e \in \mathcal{E}(\epsilon, x, \theta)} \lambda^e(\epsilon, x, \theta) = 1 \text{ and } \lambda^e(\epsilon, x, \theta) \geq 0 \right\}. \quad (9)$$

The function $\lambda(\cdot)$ is restricted to the set Λ ; it must be a valid mixing distribution between equilibria. Thus, meaningful inference can still be performed as long as the equilibrium sets actually restrict player actions to a nontrivial set of entry probabilities for at least some observed markets. Moreover, when there is a unique equilibrium with probability approaching one for markets with observable covariates x given θ , then $\lambda(\epsilon, x, \theta)$ is degenerate and the bounds for markets of type x are equality restrictions. Suppose for some θ , I observe a market type that implies a unique equilibrium for all values of ϵ .³⁰ Then, for this x , the model would imply three equality restrictions on the data with which to test θ , the same number supplied in the incomplete information case when equilibrium for a market type is unique. The more such markets are observed, the more restrictions can be collected to test the model parameters. Even if all market types exhibit multiple equilibria with high probability, nontrivial restrictions on θ can still be found because the underlying equilibrium

³⁰This would be the case if, for example, the vector of competitive effects was zero, $\theta_{i\delta} = 0$.

sets will restrict firm entry probabilities.

5 Inference

My goal is to conduct inference on the structural parameters of the model, which include the parameters of the payoff functions and the information structure. Inference is challenging because these parameters are set identified and the model includes an infinite dimensional nuisance parameter, the selection mechanism λ . In semi-parametric likelihood models such as this one, Chen et al. (2011) have proposed inference based on the profiled sieve likelihood ratio. They show that, although the distribution of the sieve profile likelihood ratio is complicated when the model is not point identified, it can be consistently estimated using the weighted bootstrap procedure. Importantly, this consistency result holds even when the rate of convergence is non-standard (as would be the case when λ is not point identified). This section shows how to apply their inference technique to the flexible information structure entry model.

Likelihood Representation of the Identified Set

Assume the econometrician observes data from M independent markets $\{x_m, y_m\}_{m=1}^M$, each of which comes from a data-generating process defined by the true parameters (θ^0, λ^0) . The true model is complete in the sense that it includes both a well defined model and a valid equilibrium selection mechanism, it maps onto a unique point in the space of outcome distributions of y given x . The partial identification problem arises because there may be multiple parameters (θ', λ') that generate the same conditional outcome distribution as (θ^0, λ^0) .

The sample log-likelihood function can be written as:

$$L_M(\theta, \lambda) = \frac{1}{M} \sum_{m=1}^M \log \Psi(y_m | x_m, \theta, \lambda). \quad (10)$$

The limit of the log-likelihood function, $L(\theta, \lambda) = E[\log \Psi(y|x, \theta, \lambda)]$, will be maximized at (θ^0, λ^0) . However, the maximizer is *not* assumed to be unique, i.e., there may exist (θ', λ') such

that $L(\theta^0, \lambda^0) = L(\theta', \lambda')$.³¹ As I am primarily interested in θ^0 , I treat the selection mechanism λ as a nuisance parameter and focus on θ as the object of interest. The identified set can be represented as the set of maximizers of L .

$$\Theta_I = \operatorname{argsup}_{\theta \in \Theta} \sup_{\lambda \in \Lambda} L(\theta, \lambda)$$

Therefore, to construct a confidence set for the identified set I will invert a test for inclusion in the set of optimal points of the likelihood function. Because only points in the identified set optimize $L(\cdot)$, this representation of the identified set is sharp.³²

Sieve Profile Likelihood

Using $L_M(\theta)$ directly for inference is infeasible because computation of $L_M(\theta)$ involves an optimization over the infinite dimensional selection mechanism and the integral within $\Psi(y|x, \theta, \lambda)$ does not have a closed form. Although semi-parametric likelihood methods such as penalized maximum likelihood could be used if the model were point identified, they have not been shown to be consistent when the model is partially identified.

The sieve profile likelihood function can be used to conduct inference on the identified set by converting an infinite dimensional problem to a finite dimensional one. First, to numerically approximate the integral over ϵ in the definition of $\Psi(y|x, \theta, \lambda)$, I evaluate the function at a finite number of R points drawn from the distribution of ϵ :

$$\Psi^R(y|x, \theta, \lambda) = \frac{1}{R} \sum_{r=1}^R \sum_{e \in \mathcal{E}(\epsilon^r, x, \theta)} \lambda^e(\epsilon^r, x, \theta) \tilde{\Psi}(y|\epsilon, x, \theta, e), \quad (11)$$

The methods of Pakes and Pollard (1989) can be used to show that Ψ^R converges to Ψ uniformly in Θ as R increases. I assume that the number of draws grows faster than the number of observations so that the impact of simulation error is asymptotically negligible.³³

³¹In the event that the maximum is unique, the model is point identified. In that case the inference procedure presented in this section is still valid.

³²Beresteanu et al. (2011) propose an alternative representation of the identified set based on the Aumann expectation. An advantage of the likelihood based approach is that it has a clear interpretation under mis-specification of inferring the set of parameters which minimize Kullback-Leibler convergence.

³³Evaluation points may be selected either through simulation using random or pseudo-random sequences or through deterministic sequences such as Halton sequences. In practice, I use Halton sequences, which have been shown to

The sieve profiled likelihood replaces Ψ with Ψ^R and replaces the infinite dimensional parameter λ with a finite dimensional sieve function. Let Λ^R represent a series of sieve spaces that approximate Λ as $R \rightarrow \infty$ (Chen, 2007). Then the sieve profiled likelihood is,

$$\hat{L}_M(\theta) = \max_{\lambda \in \Lambda^R} \frac{1}{M} \sum_{m=1}^M \log \Psi^R(y_m | x_m, \theta, \lambda). \quad (12)$$

Notice that, because the integral over ϵ is approximated (11), λ is only evaluated at a finite set of points.³⁴ Therefore, any sieve function which interpolates over these points is numerically equivalent. Although, the precise definition of the sieve space does not enter into the objective function directly, for concreteness, the sieve approximation can be a spline where the knots are the R points where λ is evaluated by Ψ^R .³⁵ Under standard regularity conditions, \hat{L}_M converges uniformly to its expectation over Θ and that the optima of \hat{L}_M converge to points within the identified set (see Chen et al., 2011, Theorem 3.1 and the remarks thereafter).

Computing \hat{L}_M involves an optimization over the set Λ^R . However, in practice this is simply a finite dimensional constrained optimization problem (the unknown variables are the weights $\lambda^e(\epsilon^r, x, \theta)$ at those ϵ^r for which there are multiple equilibria). Moreover, the objective function is concave in λ and the constraints are linear. This form of optimization problem is tractable for modern nonlinear optimization packages even with a very high-dimensional parameter space. Appendix B provides more computational detail on the optimization step.

Likelihood Ratio Test and the Weighted Bootstrap

Inference can be done via a likelihood ratio test on the statistic,

$$LR(\theta) = 2 \left(\max_{\theta' \in \Theta} \hat{L}_M(\theta') - \hat{L}_M(\theta) \right).$$

have desirable fast-convergence properties vis-a-vis pseudo-random sequences in numerical simulations. For details about Halton sequences, see Bhat (2001).

³⁴Recall that x is a discrete variable. In principle the method could be extended to allow for continuous covariates by introducing a sieve-space for (x, ϵ) , although it would quickly become intractable due to the curse of dimensionality.

³⁵If simple linear interpolation (which is a special case of a spline) is used, then the sieve approximation of λ will be bounded between 0 and 1. Higher order splines may not preserve this property, however a more complicated sieve space could be used.

Under the null hypothesis, $H_0 : \theta \in \Theta_I$,

$$\lim_{M \rightarrow \infty} P(LR(\theta) \leq c_M(\theta, 1 - \alpha)) \geq 1 - \alpha, \quad (13)$$

where $c_M(\theta, 1 - \alpha)$ is the $1 - \alpha$ quantile of the distribution of $LR(\theta)$ under the null hypothesis that $\theta = \theta^0$. If this quantile were known, (13) could serve as a test of H_0 that is consistent at level α . To gain a feasible test, $c_M(\theta, 1 - \alpha)$ must be replaced with a consistent estimator of the cutoff. The weighted bootstrap offers one method for generating this estimator.

The weighted bootstrap uses B different weighted likelihood functions, which share the same asymptotic distribution as the standard likelihood function, to approximate the distribution of the likelihood ratio test statistic. The *weighted* likelihood function is defined as a function of weights $w = (w_1, \dots, w_M)$, which are independent of the data and distributed such that $E(w_i) = 1$ and $w_i > 0$.³⁶

$$\hat{L}_M(\theta, w) = \max_{\lambda \in \Lambda(\theta)} \frac{1}{\sum w_m} \sum_{m=1}^M w_m [\log \Psi^R(y_m | x_m, \theta, \lambda)]. \quad (14)$$

The weighted likelihood ratio statistic is also defined analogously to its unweighted counterpart.

The key to the validity of the weighted bootstrap is the independence of the bootstrap weights from the data. Independence insures that the asymptotic distribution of each weighted-likelihood function is identical. Moreover, the asymptotic distribution of the standard log-likelihood is also identical to each weighted bootstrap draw. Therefore, the centered quantiles of a collection of weighted bootstrap estimators provides a consistent estimate of the quantiles of the asymptotic distribution of the sieve profiled log-likelihood under the null hypothesis.

The resulting estimator of $c_M(\theta, 1 - \alpha)$ is consistent and can be used to construct valid confidence regions for θ . Using the weighted bootstrap to calculate $\hat{c}_M(\theta, 1 - \alpha)$, I test each parameter θ for inclusion in the identified set by replacing $c_M(\theta, 1 - \alpha)$ with $\hat{c}_M(\theta, 1 - \alpha)$ in (13). Inverting this

³⁶I draw from the standard exponential distribution to produce these weights, some experimentation using the log-normal distribution produced very similar results.

test, I collect all θ which are not rejected as my confidence region for the identified set.³⁷

$$\hat{\Theta}_{CR} = \{\theta : LR(\theta) \leq \hat{c}_M(\theta, 1 - \alpha)\}$$

The set $\hat{\Theta}_{CR}$ is not an estimate of the identified set, but rather the confidence region for θ , the collection of points that are within the identified set with confidence level $1 - \alpha$. The computational details of this procedure are discussed in Appendix B.

6 Application: Supercenters and Rural Grocery Markets

The supercenter has dramatically altered the retail grocery industry and poses a new challenge to grocery retailers as a never before seen competitor format. A supercenter combines a discount store, a grocery store, and possibly several other retail services (pharmacy, tires, gas, etc.) into a single outlet of roughly 175 thousand square feet.³⁸ Several studies have examined the competitive effect of supercenters on traditional grocery stores. Hausman and Leibtag (2007) use a national panel of households to study the consumer-welfare effects of supercenters. They find that supercenters offer consumers lower prices on products and induce other grocery retailers to lower their prices, thus providing both direct and indirect positive effects on consumer welfare. Basker and Noel (2009) analyze store-level price data from 175 US cities and find that Wal-Mart's prices on average are 10 percent lower than those of its competitors and that Wal-Mart entry causes competitors to decrease their prices by 1-1.2 percent. Singh et al. (2006) investigate a supercenter's effect on grocery-store revenue. Using a frequent-shopper database from a single grocery store before and after entry by a Wal-Mart supercenter two miles away, they find that supercenter entry caused a 17 percent decline in sales revenue. Matsa (2011) shows that traditional grocery stores improve their quality of service by reducing the probability of items being out of stock in response to supercenter entry. Beresteanu et al. (2010) have investigated how supercenters affect store openings and closings of chain retail groceries competing in a wide geographic region.

³⁷In practice, I use simulated annealing to select a sample of points in the parameter set to test. This procedure is similar to the procedures of Chernozhukov et al. (2007) and Romano and Shaikh (2008) except that I use the weighted bootstrap to determine $\hat{c}_M(\theta, 1 - \alpha)$.

³⁸Singh et al. (2006) provide a full discussion of the supercenter format.

This application will focus on store openings and closings in rural areas, which have received less attention than urban markets. The impact of supercenters in rural areas is particularly important due to the possible creation of “food deserts”—areas where consumers must travel several miles to access groceries. In the popular press, the possibility of supercenters crowding out small grocery stores in rural areas has received much attention.³⁹ It is tempting to draw parallels between the decline of single-store discounters and the future of traditional supermarkets.⁴⁰ Using the methods developed above, I will examine the impact of supercenters on the profitability and viability of rural grocery store.

There are significant differences between rural grocery markets and their urban and suburban counterparts. Whereas grocery retail in denser areas is usually dominated by large chain outlets of more than 100 stores, stores in rural markets tend to be smaller and are more likely to be independent. In my sample, the average store size is 18 thousand square feet compared to the national average of 28 thousand square feet. In addition, 76 percent of stores in rural markets are not vertically integrated. These may be either local chains or single store enterprises.⁴¹ These differences suggest that key aspects of the supercenter format—large scale and a highly integrated distribution network—may be less effective in rural settings. Alternatively, the success of big-box discount stores in rural areas, particularly Wal-Mart, suggests that supercenters may induce rural residents to travel long distances, reducing demand for stores in rural markets. Hence, degree of substitutability between supercenters and rural grocers is an important empirical question.

To examine the impact of supercenters on grocery store values, I estimate a model of grocery store entry and exit that will compare the impact of a supercenter in the vicinity of a rural market on grocery store profits to the impact of a second competitor within the rural market. I assume that the decision of the supercenter to locate near the market is exogenous. This assumption seems reasonable in the rural grocery store setting because supercenters’ catchment areas are much larger

³⁹For example, in 2004 the National Trust for Historic Preservation placed the entire state of Vermont in its register of endangered sites because of Wal-Mart’s plans to begin supercenter expansion into the state. Vermont was already home to standard Wal-Mart stores (“Preservationists call Vermont Endangered, by Wal-Mart,” *New York Times*, May 25, 2004). For other examples see <http://walmartwatch.com>.

⁴⁰Jia (2008) has found that the growth of chain discount stores (Wal-Mart and Kmart) explains 40 to 50 percent of the decline in the number of small discount stores between 1978 and 1997.

⁴¹In the data many stores are listed as “independent” with no ownership data, so a precise breakdown between single store owners and small chains is not available.

than the rural markets under consideration. A supercenter’s location decision is likely to be driven by county-level demographics, road locations, and the location of its own distribution centers, rather than the degree of competition from small stores in the periphery of its catchment area. The assumption that national-level firms’ decisions are exogenous to those of local stores greatly simplifies the model and has been employed in other studies of strategic interaction between local firms (Akerberg and Gowrisankaran, 2006). I deal with the simultaneity between local stores’ entry decisions and profits by modeling the decisions of local grocers as a discrete game in which they attempt to maximize store level profits.⁴²

The application compares three specifications where only the information structure differs: assuming pure incomplete information, pure complete information, and flexible information structure models. Very few local market characteristics are observed, so there is a substantial amount of unobserved heterogeneity that is publicly known to all players but unobserved by the researcher. Firms undoubtedly condition on this information, such as road networks, city zoning laws, and grocer characteristics, when considering entry. In addition, firms have private information about their own prospects, such as managerial efficiency and opportunity costs, which generate uncertainty about each other’s actions. As I show, abstracting away from either public or private shocks will have implications for the model’s fit and the resulting counterfactual analysis.

Data

My primary data source is annual extracts from the Trade Dimension TDLinX database of all grocery-store locations in the United States from 1995 to 2006.⁴³ I supplement this data with demographic information from the 2000 decennial US census. To define markets, I assume that a local grocery store is in competition with other grocery stores located within the same Zip code. Zip codes themselves are route assignments and not geographic areas, so the US Census bureau has

⁴²This assumption abstracts away from chain effects (business-stealing and economies of scope) for those stores that are part of chains, but as the majority of stores are not vertically integrated, and because rural markets are much more thinly distributed than urban or suburban ones, chain effects should be minimal in this setting. Holmes (2011) estimates chain effects in the context of a monopoly model.

⁴³This data was graciously provided by Paul Ellickson. Various extracts from the Trade Dimensions database have been used in several empirical studies investigating retail industries (Ellickson, 2007; Beresteanu et al., 2010; Holmes, 2011; Orhun, 2013).

developed Zip Code Tabulation Areas (ZCTAs). Roughly speaking, ZCTAs map each census block to the Zip code of the majority of its residents.⁴⁴ I use the address information for each store to link that store to a year-2000 ZCTA. ZCTAs provide a reasonable approximation of a grocery–store catchment area in the markets under examination.⁴⁵

I will study the entry and exit patterns within rural grocery markets between 1998 and 2002. To be included in the dataset as a rural grocery market, a ZCTA must (i) have fewer than 15,000 people, (ii) have a population density of fewer than 750 people per square mile, (iii) have had at least one grocery store in operation between 1994 and 2006, (iv) have had no more than two grocery stores in simultaneous operation between 1994 and 2006, and (v) have no other grocery store active within 5 miles of the ZCTA in 2000. The first two restrictions focus the study on rural markets. Restrictions (iii) and (iv) on the number of stores active in the ZCTA potentially introduces an endogenous selection problem, but is necessary to maintain computational feasibility and to eliminate those ZCTAs that are not viable grocery markets (such as military sites).⁴⁶ Using a wider period (1994-2006) for selection than analysis (1998-2002) helps to reduce concern about endogenous selection. This procedure includes zip codes where no activity occurs but it is feasible that a store might open, as entry occurs in other years. It also avoids including markets where a third potential entrant may be considering entry that does not occur during the analysis period but is realized in earlier or later years. Because growing markets are likely to experience entry following the analysis period, I interpret my results as pertaining to viable and stable monopoly and duopoly markets. The final restriction (v) is meant to ensure that the markets are isolated economic units that are not heavily influenced by entry and exit outside of the market boundaries.⁴⁷

Table 1 displays summary statistics on the markets in the dataset. The expansion of supercenters is apparent from the evolution of the distance from the markets to a supercenter over time.

⁴⁴For a full description of the creation of ZCTAs, see US Census, <http://www.census.gov/geo/ZCTA/zcta.html>, (accessed April 1, 2011).

⁴⁵In their study of the loyalty card program of a grocery store located in a “small East Coast town,” Singh et al. (2006) find that the average customer lives 3.5 miles from that store, whereas 78 percent of customers live within 5 miles of the store. A five–mile radius translates roughly to a catchment area of 80 square miles. The average land area of a market as defined in this data set is 144 square miles.

⁴⁶Including markets where three stores are open simultaneously would expand the dataset by nine percent. Including these ZCTAs does not appear to affect the descriptive results presented in Section 6.

⁴⁷The set of markets does not change significantly if we use years around 2000 for this restriction.

Between 1998 and 2002, the median distance to a supercenter decreased by 10 miles. For the purposes of this application, a market is considered to be in the vicinity of a supercenter if the minimum distance to a supercenter is less than 20 miles in 2000.⁴⁸

Preliminary Analysis

For descriptive purposes, I first present simple cross-sectional evidence from the data. Table 2 presents the distribution of active firms in each market in 2002. The first two columns condition the distribution on whether (or not) there is a supercenter in the vicinity of the market. The table indicates that the presence of a supercenter is associated with fewer duopolies and more monopolies in the local market and only a very slight increase in the number of unserved markets. This is consistent with the presence of a supercenter negatively affecting traditional grocery stores in small markets. Table 3 examines the number of active stores in markets broken down by market size and whether the market is in the vicinity of a supercenter. As expected, market size is strongly correlated with the number of firms in the market. In comparison, the relationship between supercenters and the distribution of active stores is less strong.

To demonstrate how entry and exit patterns change with competition once observable market characteristics are controlled for, Tables 4 and 5 present probit regressions on the propensity of firms to enter and exit, respectively. For consistency with the results in the following section, I consider entry over a five year period.⁴⁹ These regressions do not control for the simultaneous entry and exit decisions of rival firms. Thus, the results should not be interpreted causally.⁵⁰

Column I of Tables 4 and 5 show the results of a probit model on entry and exit (respectively) controlling for only supercenter presence, log population, and presence of a competitor. Intuitively, supercenter presence is associated with less entry and more exit, but its importance relative to having a local competitor is small. Column II shows that this result is maintained when we add

⁴⁸The results are qualitatively unchanged by using 1998 or 2002 as the base year, or by extending the supercenter radius to 25 miles. I have also experimented with using two distance bands of 15 and 30 miles and found qualitatively similar results.

⁴⁹To be specific, entry is defined as a firm who was not present in 1998 being present in 2002, and exit is a firm who was present in 1998 being absent in 2002.

⁵⁰It is interesting to note that if one were to attempt to estimate a dynamic model of incomplete information using these data (e.g., Aguirregabiria and Mira, 2007; Bajari et al., 2007), these regressions would be similar to the “first stage” estimates of the policy function.

a richer set of controls involving regional differences, local household income, and whether the supercenter is a new opening itself. None of these controls are statistically significant and I exclude them from the general model for reasons of parsimony.

Tables 4 and 5 include log population as a control in columns I and II. In the full model, I discretize population in order to avoid adding an additional dimension to the sieve space for the selection mechanism. The discretization also allows us to examine possible non-linearities in the effect of (log) population. The specification includes dummies for whether the population is larger than 3,000 or 6,000.⁵¹ The cutoffs for the dummies were chosen to approximate the one-third and two-thirds quantiles of the data. Columns III and IV of Tables 4 and 5 replace log population with this discretization. The discretization has little effect on the estimated coefficient of supercenter presence or the existence of local competition. Using the dummies, we see that the relationship between entry (and exit) and population is non-monotonic, but these results should not be given a causal interpretation. This result suggests that there is indeed some endogeneity affecting these preliminary results.

Finally, Column V of Tables 4 and 5 adds an interaction term between the presence of a competitor and population. It appears that the degree of competition between firms varies with market size. Therefore, I include an interaction between the competition effect and market size in the structural model.

Bresnahan-Reiss Symmetric Firms Complete Information Model

Before turning to the full model in the following section, it is instructive to present the results of the same data using a simpler specification of Bresnahan and Reiss (1991b). This model is nested within the model of the following section with three major restrictions: There is no private information shock, unobserved heterogeneity is perfectly correlated, and firms are homogeneous, so

⁵¹Both coefficients should be added to compute the total effect of population when the market is larger than 6,000 persons. Probit regressions such as those in Tables 4 and 5 are robust to slight changes in the population cutoffs.

observed heterogeneity must be ignored. The latent profit function becomes,

$$\pi_i(y_i, y_{-i}; x_m, \theta) = \begin{cases} \mu(x_m) + y_{-i}\delta(x_m) + \epsilon_m & \text{if } y_i = 1 \\ 0 & \text{if } y_i = 0 \end{cases}. \quad (15)$$

where δ and μ are linear functions of x_m , which contains only market specific (not firm specific) observable characteristics. Although this model admits multiple equilibria, Bresnahan and Reiss (1991b) point out that it generates a unique prediction on the *number* of firms that enter, as opposed to their identities. This allows the model to be estimated as an ordered probit where the dependent variable is the number of firms in the market.

Table 6 presents two specifications of the Bresnahan-Reiss model. Column I includes only a constant for the competition effect. Column II allows for interactions between market characteristics and the competition effect. Comparing these results to the probits in Tables 4 and 5, the coefficients on the population dummies are much more intuitive. This may be evidence of the importance of controlling for endogeneity of rival decisions in a structural manner. Second, we see that the impact of competition from local rivals appears to be much more severe than competition supercenters, although both negatively affect profits. This relationship will also appear in the results of the full model below. Finally, comparing the two columns, it appears that the competition affect does seem to vary with market characteristics. It appears that the presence of a supercenter significantly hardens competition between rivals, whereas rivalry is softer in the highest population markets. Therefore, we allow the competition effect to vary with market characteristics in the following section.

Because the Bresnahan-Reiss specification is nested within the full model of the following section, it is easy to test the results of Table 6 against the full model. As the following section will show, the restrictions of this model are rejected, which is not surprising given the importance of firm entry costs to profitability (these must be excluded from the Bresnahan-Reiss model because it assumes firms are homogeneous). Overall, these results provide some indication that the restrictions needed to avoid the problem of multiple equilibria in entry games are not innocuous.

7 Results from the Flexible Information Model

I now apply the full model described in Section 2 to data on entry and exit patterns of rural grocery stores. I model the decision to operate over a five year period as a one shot simultaneous decision where store managers receive profit from running the store and any sell-off value from the store as a lump sum if they chose to operate. State variables at the market level are whether or not a supercenter is present within 20 miles in 2000 and dummies for a population greater than 3,000 or 6,000 people. In principle, the effect of population could be modeled as a linear term, but I choose not to do so given the non-monotonic response to population found in the preliminary estimates (see Tables 4 and 5).⁵² Firm level state variables are whether or not the firm was operating in 1998, which indicates whether the firm must sink investment in order to begin operation. The outcome variables are whether or not each firm was operating in 2002. I assume that opening a store affects only a firm's costs and not demand, so it can be used as an exclusion restriction. That is, the existence of a rival firm in 1998 only affects firm i 's entry decision through its impact on i 's expectation that he will compete in a duopoly.⁵³ The model requires the strong assumption that shocks for the game are uncorrelated with the determinants of the market structure prior to 1998.⁵⁴ Profits of firm i conditional on entry can be written as,

$$\pi_i(x_{im}, y_{-im}) = \mu(x_{im}) + y_{-im}\delta(x_{im}) + \nu_{im} + \epsilon_{im}$$

Where $\mu(x_{im})$ is a function defining baseline monopoly profits and $\delta(x_{im})$ is a function defining the competition effect.

⁵²The discretization of population is computationally convenient for two reasons. First, discretization allows me to solve the model once for many observations, speeding up computation of the likelihood. Second, if x were continuous, the selection mechanism would have to be modeled with a $K + 2$ dimensional sieve where K is the number of continuous dimensions of x .

⁵³The exclusion restriction assumption is needed to provide identifying power; the methodology used here could be applied without the exclusion restriction, but the identified set would expand accordingly. This suggests a robustness test for the exclusion restriction assumption.

⁵⁴This assumption will be violated if public shocks are persistent over time. However, the assumption that errors are uncorrelated across time is common in empirical applications of games. Relaxing this assumption within a dynamic setting is an avenue for future work.

The baseline monopoly profit function is

$$\begin{aligned} \mu(x_{im}) = & \mu_0 + \mu_1 \mathbf{1}[Pop_m > 3k] + \mu_2 \mathbf{1}[Pop_m > 6k] \\ & + \mu_3 \mathbf{1}[Supercenter_m < 20mi] - \mu_4 \mathbf{1}[iInactive \text{ in } 1998]. \end{aligned}$$

I will refer to μ_4 as the entry cost; it measures the extra costs of opening a new grocery store.⁵⁵

The effect of competition within the local market is captured by $\delta(\cdot)$ which is also assumed to be linear:

$$\delta(x_{im}) = \delta_0 + \delta_1 \mathbf{1}[Pop_m > 3k] + \delta_2 \mathbf{1}[Pop_m > 6k] + \delta_3 \mathbf{1}[Supercenter_m < 20mi].$$

This specification is in line with the model of Bresnahan and Reiss (1990) in which competition can affect both variable profits and fixed costs (through entry barriers). The main parameters of interest are the effect of supercenters on local grocery store profits and the competitive effect of other local grocery stores on store profits.

Confidence Set for the Structural Parameters

For comparison purposes, I estimate the model under the incomplete information and complete information assumptions that are commonly employed in the entry literature. For the complete and incomplete information models, I impose point identifying assumptions and use traditional inference methods.⁵⁶ For the full model, I do not assume point identification and instead employ the inference techniques described in Section 5. These techniques do not provide point estimates. Instead, I report the 95 percent confidence sets for the identified set and a 95 percent confidence region for the true parameter. The first covers the entire identified set at a confidence level of 95 percent, whereas the second (which is a subset of the first) covers the true parameter values with

⁵⁵The effect on profits from opening a store are assumed to be due to construction and start-up costs, rather than low demand for a new store when a rival is active.

⁵⁶Specifically, for the complete information model, I assume that, when multiple equilibria are present, one of the pure strategy equilibria is chosen with equal probability (the results are similar if I assume the equilibrium which maximizes total profits is chosen); for the incomplete information model, I assume the data are generated by a single equilibrium.

a confidence level of 95 percent.

Table 7 presents the results.⁵⁷ The full model nests both the complete and the incomplete information frameworks. Therefore, we can use the full model to test the two restricted models by checking to see whether any element in the confidence set of the full model satisfies the conditions $\sigma_\epsilon^2 = 0$ (incomplete information) or $\sigma_\epsilon^2 = 1$ (complete information). As shown in Table 7, the incomplete information model is rejected at the 0.05 level, whereas the complete information model cannot be rejected. On the other hand, the model does not reject $\sigma_\epsilon^2 < .5$, so more than half the variance may be generated by a private error. Thus there is no empirical evidence for focusing exclusively on the pure complete information model.

As in most discrete-choice models, it is difficult to interpret the parameter confidence intervals. The counterfactual calculations presented in the next section clarify the implications of the model. Nonetheless, the results indicate that the presence of a supercenter has a mild negative effect on the value of a grocery store, although this is not statistically significant. The effect of local competition appears to be stronger. The baseline effect of population on monopoly profits is monotonically increasing (in contrast to the preliminary results of Tables 4 and 5, which did not control for endogeneity of firm decisions), whereas population's effect on competition is ambiguous. These results are broadly in line with the findings of the Bresnahan-Reiss style model in Table 6.⁵⁸ In contrast, the entry cost, which is not accounted for in Table 6 is large and positive. Overall, the restrictions of the model in Table 6 are rejected.

Counterfactual Experiments

Using the confidence region for the identifiable parameter, I construct bounds for counterfactual statistics, such as the change in firm value from changes in the market structure, or the expected number of grocery stores in markets of different sizes. These counterfactuals are functions of both the parameters of the model, θ , and the selection mechanism, λ (i.e., a counterfactual is some

⁵⁷Because the inference procedure for the full model yields a joint confidence region, I report projections of this region onto parameter axes. For this reason, Table 7 exaggerates the size of the confidence sets of the full model. Many parameter values within the cartesian product of these intervals are outside the confidence set. The counterfactuals in the next section operate using the true confidence region, a subset of the "box" reported in Table 7.

⁵⁸Both models are scale normalized, so the relevant comparison is in ratios of coefficients rather than their absolute value.

function $f(\theta, \lambda)$). I have derived a confidence region for θ but not λ . A very conservative method for deriving bounds would be to pair all $\theta \in \hat{\Theta}_{CR}$ with *any* valid selection mechanism to form the range of values in the counterfactual experiments.

$$\left[\inf_{\theta \in \hat{\Theta}_{CR}} \inf_{\lambda \in \Lambda(\theta)} f(\theta, \lambda), \sup_{\theta \in \hat{\Theta}_{CR}} \sup_{\lambda \in \Lambda(\theta)} f(\theta, \lambda) \right] \quad (16)$$

This method yields extremely wide confidence intervals, and is unappealing because it admits selection mechanisms which clearly do not fit the observed data. Moreover, because I can compute the likelihood, $L_M(\theta, \lambda)$, I know that many (θ, λ) pairs with $\theta \in \hat{\Theta}_{CR}$ have a likelihood that is far from the optimum. If λ were finite dimensional, inference on λ would make use of an inverted likelihood ratio test, i.e.,⁵⁹

$$\hat{\Lambda}(\theta) = \left\{ \lambda \in \Lambda(\theta) : 2 \left(\sup_{\theta' \in \Theta} L_M(\theta') - L_M(\theta, \lambda) \right) \leq \hat{\kappa}_M \right\}$$

Where $\hat{\kappa}_M$ is a consistent estimate of κ_M , the $1 - \alpha$ quantile of the distribution of the likelihood ratio. Unfortunately, there is no known procedure to consistently approximate κ_M when λ is infinite dimensional. Instead, I follow the following procedure to restrict the set of plausible selection mechanisms for each $\theta \in \hat{\Theta}_{CR}$. The procedure stems from the observation that *all* (θ, λ) such that θ that are not in the coverage set of Θ_I (i.e., the set presented in Column 3 of Table 7), or equivalently, those for which the *profiled* likelihood ratio is above \hat{c}_M , can be rejected with 95 percent confidence.⁶⁰ Intuitively, two parameterizations with the same likelihood are equally likely, so I propose to use $\gamma_M \hat{c}_M$, where $\gamma_M > 1$ as the basis for an ad hoc cutoff of for restricting the set of selection mechanisms λ to pair with $\theta \in \hat{\Theta}_{CR}$ when creating counterfactual confidence bounds. This procedure will be consistent, but conservative, as long as $\gamma_M \hat{c}_M \geq \kappa_M$ holds as $M \rightarrow \infty$. One way of interpreting this cutoff is that, for (θ, λ) with $\theta \in \hat{\Theta}_{CR}$, λ is excluded as a possible selection mechanism only if θ would have been comfortably excluded from the confidence region if λ were

⁵⁹Recall that $L_M(\theta)$ is the profiled likelihood which already optimizes over λ . Of course, if λ were finite dimension, inference on θ and λ could be performed jointly.

⁶⁰See Appendix B.3 for details on the calculation of \hat{c}_M .

its only valid selection mechanism. Therefore the bounds I present are,

$$\left[\inf_{\theta \in \hat{\Theta}_{CR}} \inf_{\lambda \in \bar{\Lambda}(\theta)} f(\theta, \lambda), \sup_{\theta \in \hat{\Theta}_{CR}} \sup_{\lambda \in \bar{\Lambda}(\theta)} f(\theta, \lambda) \right],$$

where $\bar{\Lambda}(\theta) = \{\lambda \in \Lambda(\theta) : 2(\sup_{\theta' \in \Theta} L_M(\theta') - L_M(\theta, \lambda)) \leq \gamma_M \hat{c}_M\}$. In practice I use $\gamma_M = 1.5$, but in experimentation I have found the results are very insensitive to changes in γ_M .⁶¹

Effects on Firm Valuations

This section investigates the effect of market structure on firm valuations. I use these counterfactuals to compare the results of the full model with those of the more restrictive complete information model. The bounds for the incomplete information model are not included as that model is rejected by the full model (effectively, its confidence set is empty).⁶² Firm valuations are the expected payoff from operating the firm before shocks are revealed and the equilibrium is selected. The formula for the expected value of the firm at the start of the period is:

$$E[\pi_i(x; y) | x; \theta, \lambda] = \int \sum_{e \in \mathcal{E}(x, \epsilon)} \lambda^e(x, \epsilon) \left(\varrho_i^e(x, \epsilon; \theta) (\mu(x) - \varrho_{-i}^e(x, \epsilon) \delta(x) + E[\nu_i | \nu_i \geq \chi_i^e(x, \epsilon)] + \epsilon_i) \right) dF(\epsilon). \quad (17)$$

The object of interest is the relative change in firm value of moving a firm from state x to state x' , i.e., $\frac{E[\pi_i(x'; y)] - E[\pi_i(x; y)]}{E[\pi_i(x; y)]}$.

First, I examine the effect of adding a supercenter in the vicinity of a market where none existed. Bounds construct using the procedure describe above are presented in Table 8. The bounds for the full model indicate that supercenters may decrease expected long-run firm profits by up to 24.6 percent but may also generate up to a 9.3 percent increase in profits.

The difference between the bounds of the complete information model and the full model are substantial. First, consider the differences in upper bounds. The upper bound under the complete

⁶¹Of course, I cannot check all valid selection mechanisms. Instead I sample from the set of valid selection mechanisms within the sieve by (i) perturbing the selection mechanism which optimizes $L_M(\theta)$ (as well as including the unperturbed optimizer) (ii) drawing uniformly from the selection mechanism and (iii) testing the selection mechanisms that generate the extreme points in (7).

⁶²The confidence bound under the null that the model is correctly specified are available from the author.

information model is substantially lower for all three market sizes. The difference is a result of relaxing assumptions about the information structure. In the complete information framework, firms can condition on their rival's public shock and avoid negative outcomes, so the benefits from less rival activity only appear on the margin. On the other hand, if the firms are uncertain about rival entry, a reduction in entry probabilities reduces the chance of simultaneous entry resulting in negative payoffs. Lowering the probability of negative profit outcomes can substantially benefit firms by alleviating their coordination problem. The complete information model assumes away this coordination problem, whereas the full model recognizes this potential benefit to supercenter entry. In other words, because negative profit outcomes never occur under complete information assumptions, the complete information model is unable to capture the full benefit of reducing rival entry when firms are uncertain.

Furthermore, in duopoly markets, the lower bound of the supercenter effect for the full model is substantially lower than the bound calculated using complete information assumptions. In the complete information case, firms are able to avoid negative profit outcomes, so the decrease in profits is due to both fewer profitable opportunities and less-positive profits given those opportunities. The full model includes the possibility that the harsher environment resulting from supercenter entry makes profits even more negative in the event of ex-post regret. This result is particularly stark if firms begin the period as a duopoly, when both firms are more likely to choose to operate.

In sum, the pure strategy complete information model abstracts away from uncertainty about opponents. This ignores the coordination problem that firms face when entry by both will result in negative payoffs.⁶³ Allowing for uncertainty between firms systematically widens the bounds on the effect of supercenter entry.

Table 9 bounds the effect on firm value for a monopolist who experiences entry by a local rival. In contrast to the effect of supercenter entry, this effect is unambiguously negative under both the complete information framework and the full model. (This was not clear from the results reported in Table 7 and gives an indication as to why the counterfactual results may be more useful

⁶³The coordination problem can arise due to the play of mixed strategies; however, one could argue that mixed strategies would be avoided in this case because they result in lower expected profits for both players. Moreover, most empirical studies using the complete information model rule out mixed strategies by assumption.

than the parameter bounds themselves.) However, the lower bound is much more severe for the full model: a decrease in expected profit in the range of 46-60 percent compared to only a 23-28 percent loss when assuming complete information. The difference is again attributable to the fact that firms with complete information are able to completely avoid situations that result in negative profits. This also illustrates that relaxing the information restriction does not simply widen bounds symmetrically.

Tables 8 and 9 give the impression that entry by a local grocery store is more harmful than non-local supercenter entry. However, the bounds overlap and do not rule out the possibility that monopolists may prefer facing local competition to competition from a supercenter. I use the model to examine this question directly by computing bounds for the following statistic:

$$C \equiv \frac{E[\pi_i | \text{Monopoly, Supercenter}] - E[\pi_i | \text{Duopoly, NoSupercenter}]}{E[\pi_i | \text{Monopoly, NoSupercenter}]}.$$

The sign of C indicates whether a firm prefers to face a supercenter ($C > 0$) or local competition ($C < 0$); its magnitude measures the strength of preference scaled by monopolist firm value. Bounds for C are presented in Table 10. For the full model, the sign is technically ambiguous; however, results are again heavily tilted towards preferring supercenter competition. If stores do prefer local competition, the difference in profits is small I interpret this result as weak evidence that differentiation on the basis of location and store type is effective at blunting the cost advantages of supercenters over local grocery stores. Notably, if we impose the assumption of complete information, the bounds are much tighter and indicate a significant preference for supercenter competition over local competition.

To summarize, in this section I have compared bounds on counterfactuals under two nested assumptions. The stricter assumption—pure strategy complete information—does plausibly fit the data, but abstracts away from the possibility of uncertainty between firms. The weaker assumption of the full model allows for uncertainty between firms. The intervals generated by the full model are qualitatively different from those of the complete information model, and the reasons for these differences is linked to the elimination of private information from the model.

Availability of Grocery Stores

The previous subsection analyzed the impact of supercenter entry on the prospects of local firms. This subsection presents the implications of supercenters on the number of local grocery stores available to consumers. I focus exclusively on the full model, the results of the complete information model are broadly consistent with these results.

For the purposes of this section, I assume that exogenous market characteristics, including the presence of a supercenter, are constant over time, and that firms must pay the entry cost to be active in period t only if they were inactive in the previous period.⁶⁴ Furthermore, I make the strong assumption that both public and private shocks are iid across five year periods. With this assumption, the structural model produces a Markov chain that governs transitions over the number of stores in the market. Table 11 presents the confidence bounds on the stationary distribution of this Markov chain by market type using the full model.⁶⁵ Across market sizes, it appears that the presence of a supercenter shifts up the bounds on the proportion of unserved markets, and shifts down the bounds on the proportion of markets served by two local firms. Although the bounds are wide, the effect of supercenters on the long-run distribution of local grocery stores appears mild.

Comparing these results with the observed distribution of stores across markets (Table 3), I find that the 2000 distribution of stores is well within the bounds of the steady state distribution predicted by the model. Note that this need not be the case: the results reported in Table 11 are derived from entry and exit patterns, whereas those in Table 3 come from the static distribution of stores. This result is consistent with the view that a major shift in the availability of local grocery stores is not underway.

Some policymakers have expressed concern that supercenters may be causing a large increase in the number of markets that are unserved by local grocery stores, spawning “food deserts.” Table 11 provides upper bounds on the proportion of unserved markets. Market size is a much more important determinant of unserved markets than the presence of a supercenter, and the majority

⁶⁴This assumption is not needed for estimation. I estimate firms’ expectations of their long-run profits based on their entry and exit actions. Firms’ expectations of how the exogenous variables will change in the future is accounted for in their expectations of long-run profits.

⁶⁵The bounds from the complete information model are not presented in this section as they do not provide any additional insights beyond the results of the full model.

of markets will be served by at least one grocery store in large and medium markets. In small markets, where the proportion of unserved markets is already high (Table 3), the proportion of unserved markets can only be bound below two-thirds. It appears that the smallest markets are barely able to meet the minimum scale for even a single grocery store. The presence of a supercenter exacerbates the problem, but restrictions on supercenters would not be likely to remove the threat of small markets becoming “food deserts.”

Table 12 presents bounds on the expected number of stores in each market type as well as bounds on the “supercenter effect”—the difference in the expected number of stores (in percentage terms) in similar sized markets with and without a supercenter. These results indicate a downward shift in the bounds on the number of local stores available to consumers as a result of a nearby supercenter, although the bounds overlap. The interval for the supercenter effect is mostly negative. This echoes the descriptive results from Section 6 and other indicators in this section that the effect of supercenter entry mildly reduces the number of local stores. However, it seems unlikely that this decrease in the expected number of local-grocery-store options offsets the benefits from adding the supercenter option to consumers’ choice sets. Of course, a full analysis of consumer welfare is outside the scope of this paper.

8 Conclusion

Earlier studies of discrete games have assumed that unobserved factors of the game are either publicly observed by all players (complete information) or privately observed by individual players (incomplete information). I provide a more general model that nests these assumptions and parameterizes the extent to which unobservable components of a firm’s profits are publicly known. Using this model, information in the data can be used to make inference on the extent to which variation in firm actions is due to public or private information. The usual assumptions made by both the pure complete and pure incomplete information frameworks are testable. By using inference techniques that avoid point identification assumptions, I construct bounds on model statistics without imposing ad hoc assumptions on the information structure or equilibrium selection.

I apply the model to examine the effect of supercenters on rural grocery markets. The growth

of the supercenter format has led industry observers to inquire how supercenters alter the grocery market, and whether they are likely to displace local grocery stores. For grocery store owners, I find that entry by a supercenter is far less detrimental than entry by a local competitor, and that if supercenter entry is effective at suppressing the probability of entry by a local challenger, it may actually increase long-run profits for incumbents. This outcome is the product of a reasonable economic model that fits the data well, but is ruled out by assumption within the complete information framework.

The empirical results show that placing strong assumptions on the information structure of a game has real consequences. The incomplete information framework is rejected when tested against the general model. Although the complete information framework is not rejected, bounds produced using this framework are driven by assumptions on the information structure. This could lead the researcher to make overly strong conclusions about policy relevant statistics of interest. By incorporating both public and private information into a single model and using partial identification inference techniques, it is possible to develop meaningful confidence bounds for many statistics of interest under relatively weak informational assumptions.

References

- Akerberg, D. A. and G. Gowrisankaran (2006). Quantifying equilibrium network externalities in the ach banking industry. *The RAND Journal of Economics* 37(3), 738–761.
- Aguirregabiria, V. and P. Mira (2007). Sequential estimation of dynamic discrete games. *Econometrica* 75(1), 1–53.
- Aradillas-Lopez, A. (2010). Semiparametric estimation of a simultaneous game with incomplete information. *Journal of Econometrics* 157, 409–431.
- Augereau, A., S. Greenstein, and M. Rysman (2006). Coordination versus differentiation in a standards war: 56k modems. *The RAND Journal of Economics* 37(4), 887–909.
- Bajari, P., C. L. Benkard, and J. Levin (2007). Estimating dynamic models of imperfect competition. *Econometrica* 75(5), 1331–1370.
- Bajari, P., J. Hahn, H. Hong, and G. Ridder (2011). A note on semiparametric estimation of finite mixtures of discrete choice models with application to game theoretic models. *International Economic Review* 52(3), 807–824.
- Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010a). Computing equilibria in static games of incomplete information using the all-solution homotopy. Minnesota, Stanford, SF Fed, and UC Berkeley.
- Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010b). Estimating Static Models of Strategic Interactions. *Journal of Business & Economic Statistics* 28(4), 469–482.
- Bajari, P., H. Hong, and S. Ryan (2010). Identification and Estimation of a Discrete Game of Complete Information. *Econometrica* 78(5), 1529–1568.
- Basker, E. and M. D. Noel (2009). The evolving food chain: Competitive effects of wal-mart’s entry into the supermarket industry. *Journal of Economics and Management Strategy* 18(4), 977–1009.
- Beresteanu, A., P. B. Ellickson, and S. Misra (2010). The dynamics of retail oligopoly. University of Pittsburgh and University of Rochester.
- Beresteanu, A., I. Molchanov, and F. Molinari (2008). Sharp identification regions in games. CEMMAP Working Paper CWP 15/08.
- Beresteanu, A., I. Molchanov, and F. Molinari (2011). Sharp identification regions in models with convex moment predictions. *Econometrica* 79(6), 1785–1821.
- Berry, S. (1992). Estimation of a model of entry in the airline industry. *Econometrica* 60(4), 889–917.
- Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica* 63(4), 841–890.
- Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological* 35(7), 677 – 693.

- Bjorn, P. and Q. Vuong (1984). Simultaneous equations models for dummy endogenous variables: A game theoretic formulation with an application to labor force participation. California Institute of Technology.
- Bresnahan, T. F. and P. Reiss (1991a). Empirical models of discrete games. *Journal of Econometrics* 48(1), 57–81.
- Bresnahan, T. F. and P. Reiss (1991b). Entry and competition in concentrated markets. *Journal of Political Economy* 99, 977–1009.
- Bresnahan, T. F. and P. C. Reiss (1990). Entry in monopoly markets. *The Review of Economic Studies* 57(4), 531–553.
- Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 6B, Chapter 74, pp. 5549–5632. Amsterdam: Elsevier Science B.V.
- Chen, X., E. Tamer, and A. Torgovitsky (2011). Sensitivity analysis in semiparametric likelihood models. Yale University and Northwestern University.
- Chernozhukov, V., H. Hong, and E. Tamer (2007, September). Estimation and confidence regions for parameter sets in econometric models. *Econometrica* 75(6), 1243–1285.
- Chiou, L. (2009, Summer). Empirical analysis of competition between wal-mart and other retail channels. *Journal of Economics and Management Strategy* 18(2), 285–322.
- Ciliberto, F. and E. Tamer (2009). Market structure and multiple equilibria in airline markets. *Econometrica* 77(6), 1791–1828.
- Davis, P. (2006, September). Estimation of quantity games in the presence of indivisibilities and heterogeneous firms. *Journal of Econometrics* 134(1), 187–214.
- Ellickson, P. and P. L. E. Grieco (2013). Density versus differentiation: The impact of wal-mart on the grocery industry. *Journal of Urban Economics* 75, 1–14.
- Ellickson, P. B. (2007, March). Does sutton apply to supermarkets? *RAND Journal of Economics* 38(1), 43–59.
- Galichon, A. and M. Henry (2011). Set identification in models with multiple equilibria. *Review of Economic Studies* 78(4), 1264–1298.
- Haltiwanger, J. C., R. S. Jarmin, and C. J. Krizan (2010). Mom-and-pop meet big box: Compliments or substitutes? *Journal of Urban Economics* 67(1), 116–134.
- Harsanyi, J. C. (1973). Games with randomly distributed payoffs: A new rationale for mixed strategy equilibria. *International Journal of Game Theory* 2(1), 1–23.
- Hausman, J. and E. Leibtag (2007, December). Consumer benefits from increased competition in shopping outlets: Measuring the effect of wal-mart. *Journal of Applied Econometrics* 22(7), 1157–1177.

- Holmes, T. J. (2011). The diffusion of wal-mart and the economies of density. *Econometrica* 79(1), 253–302.
- Honoré, B. E. and E. Tamer (2006). Bounds on parameters in panel dynamic discrete choice models. *Econometrica* 74(3), 611–629.
- Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica* 76(6), 1263–1316.
- Liu, N., Q. Vuong, and H. Xu (2012). Rationalization and nonparametric identification of discrete games with correlated types. Penn State, NYU, and UT-Austin.
- Matsa, D. A. (2011). Competition and product quality in the supermarket industry. *Quarterly Journal of Economics* 126(3), 1539–1591.
- Mazzeo, M. (2002). Product choice and oligopoly market structure. *RAND Journal of Economics* 33(2), 1–22.
- Milgrom, P. R. and R. J. Weber (1985). Distributional strategies for games with incomplete information. *Mathematics of Operations Research* 10(4), 619–632.
- Morris, S. and H. S. Shin (2000). Rethinking multiple equilibria in macroeconomic modeling. *NBER Macroeconomics Annual* 15, 139–161.
- Orhun, A. Y. (2013). Spatial differentiation in the supermarket industry. *Quantitative Marketing and Economics* 11(1), 3–37.
- Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. *Econometrica* 57(5), 1027–1057.
- Paula, A. d. and X. Tang (2012). Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica* 80(1), 143–172.
- Romano, J. P. and A. M. Shaikh (2008). Inference for identifiable parameters in partially identified econometric models. *Journal of Statistical Planning and Inference* 138(9), 2786–2807.
- Seim, K. (2006, Autumn). An empirical model of firm entry with endogenous product-type choices. *RAND Journal of Economics* 37(3), 619–640.
- Shi, X. (2011). Model selection tests for nonnested moment inequality models. University of Wisconsin-Madison.
- Singh, V. P., K. T. Hansen, and R. C. Blattberg (2006). Market entry and consumer behavior: An investigation of a wal-mart supercenter. *Marketing Science* 25(5), 457–476.
- Sweeting, A. (2009). The strategic timing of radio commercials: An empirical analysis using multiple equilibria. *RAND Journal of Economics* 40(4), 710–742.
- Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies* 70(1), 147–165.

Vitorino, M. (2012). Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry. *Journal of Marketing Research* 49(2), 175–191.

Wan, Y. and H. Xu (2012). Semiparametric estimation of binary decision games of incomplete information with correlated private signals. University of Toronto and University of Texas-Austin.

A Proofs

Theorem 1 If Assumptions 1 and 2 hold, then the parameters of the payoff function $(\theta_{i\mu}, \theta_{i\delta})$ are point identified.

Proof. The proof is similar to Theorem 1 of Tamer (2003). Without loss of generality consider player 1's action, the argument for player 2 is symmetric. Player 1's best response function is,

$$\chi_i(\chi_{-i}, \epsilon, x) = -(x_i\theta_{i\mu} + \epsilon_i) + \rho_2(\chi_{-i}, \epsilon, x)\theta_{i\delta} \quad (18)$$

Where ρ_2 is player i 's rational belief about the rival firm's probability of entry based on its given strategy χ_{-i} . This probability is derived according to (2).

Assume without loss of generality that $\theta_{2\mu}^* > 0$.⁶⁶ Then $\lim_{x_2^* \rightarrow -\infty} P(y_2 = 1|x, \epsilon) = \lim_{x_2^* \rightarrow -\infty} P(y_i = 1|x) = 0$, because from (18), $\chi_2 \rightarrow \infty$ as $x_2^* \rightarrow \infty$ when holding all other parameters fixed regardless of firm 1's action (because the $|\theta_{2\mu}^*| > |\theta_{2\delta}^*|$ by Assumption 2). By the Bayesian Nash equilibrium assumption, $\rho_2(\infty, \epsilon; \theta) = 0$, the probability of firm 1 entering is,

$$\lim_{x_2^* \rightarrow -\infty} P(y_1 = 1|x) = E[P(\nu \geq \chi_1(\chi_2, \epsilon, x))] = P(\epsilon_i + \nu_i \geq -(x_1\theta_{1\mu})).$$

By our scale assumption on the distribution of $\epsilon_i + \nu_i$, this is a linear probit model, so the vector $\theta_{i\mu}$ is identified.⁶⁷ To identify the parameters of $\theta_{i\delta}$, note that, $\lim_{x_2^* \rightarrow \infty} P(y_2 = 1|x, \epsilon) = \lim_{x_i^* \rightarrow \infty} P(y_2 = 1|x) = 1$, so we have

$$\lim_{x_2^* \rightarrow \infty} P(y_1 = 1|x) = P(\epsilon_i + \nu_i \geq -(x_1(\theta_{1\mu} + \theta_{1\delta})))$$

As $\theta_{1\mu}$ is already identified by the preceding argument, we treat these parameters as known, the result is a linear probit model with a constant adjustment, so $\theta_{i\delta}$ is identified as well. \square

Theorem 2 If Assumptions 1 through 3 hold, the assumptions of the pure strategies complete information framework are testable.

Proof. The complete information assumption fixes σ_ϵ^2 at 1, so the only remaining parameters to identify are ρ and $\lambda(\cdot)$. Given the pure strategies assumption, player 1 knows y_2 with certainty when making his own entry decision (and vice versa). Given this, if we observe either both firms entering or neither firms entering, we can infer that the strategy was generated by a model with a unique equilibrium.⁶⁸

$$P(y_1 = 1, y_2 = 1, x; \theta) = \int_{\epsilon_1 = -(x_1\theta_{1\mu} + \theta_{1\delta})}^{\infty} \int_{\epsilon_2 = -(x_2\theta_{2\mu} + \theta_{1\delta})}^{\infty} dF(\epsilon_1, \epsilon_2; \rho)$$

Because the parameters of the objective function are identified by Theorem 1, ρ is the only free parameter on the right hand side. Because this expression is monotonically increasing in ρ , ρ is identified by observing the left hand side for a single market. After identifying ρ from a single market, $P(y_1 = 1, y_2 = 1|x)$ is known for all x , as it is independent of the selection mechanism.

⁶⁶The same argument with signs reversed can be carried out if $\theta_{2\mu}^* < 0$, and Assumption 2 guarantees $\theta_{2\mu}^* \neq 0$.

⁶⁷The argument is symmetric for player 2.

⁶⁸This argument assumes that $\theta_{i\delta} < 0$ for both firms, a similar argument holds with the sign reversed.

Therefore can test the assumptions of the pure strategy complete information model by checking to see whether $P(y_1 = 1, y_2 = 1|x)$ implied by the pure strategies complete information model is consistent with the observed distribution across all markets. \square

Theorem 3 If Assumptions 1 and 2 hold, the assumptions of the pure incomplete information framework are testable.

Proof. Under the assumptions for the incomplete information model, $\sigma_\epsilon = 0$ by assumption, so ρ drops out of the model, and $\epsilon_i = 0$ with probability one. As $(\theta_{i\mu}^*, \theta_{i\mu}, \theta_{i\delta})$ are point identified according to Theorem 1, $\mathcal{E}(0, x, \theta)$ is identified for every x . We can use the equilibrium set to construct constraints on $P(y|x)$ for all x , if $\#\mathcal{E}(0, x, \theta) = 1$ these are equality constraints, if $\#\mathcal{E}(0, x, \theta) > 1$ they are inequality constraints. The inequality constraints are non-trivial because under the incomplete information assumption firms take both actions with positive probability for all x . We can test the model by verifying these restrictions hold for all x . \square

Theorem 4 Given assumptions 1 and 4, the sharp identified set of θ for the incomplete information model is,

$$\Theta_I = \left\{ \theta \in \Theta : \forall y, x, \exists \tilde{\lambda} \in [0, 1]^{\bar{E}} \text{ s.t. } \begin{array}{l} P(y|x) = \Psi(y|x, \theta, \tilde{\lambda}), \\ \sum_{e \in \mathcal{E}(x; \theta)} \tilde{\lambda}^e = 1 \end{array} \right\}$$

Where \bar{E} is a constant which represents the largest possible number of equilibria the model admits for any $x \in X$ and any $\theta \in \Theta$.

Proof. We have assumed that $P(y|x)$ is observed for identification purposes via Assumption 1. For any θ , we treat λ as a restricted nuisance parameter. Let \mathcal{P} be the space of possible conditional probability distributions, as x and y are discrete, \mathcal{P} is a vector space. Given θ we can construct a set in \mathcal{P} of outcome distributions that are consistent with the model.

$$\mathcal{P}(\theta) = \left\{ \begin{array}{l} P(\cdot|\cdot) \in \mathcal{P} : \forall x, y, \exists \tilde{\lambda} \in [0, 1]^{\bar{E}} \text{ s.t.:} \\ P(y|x) = \Psi(y|x, \theta, \tilde{\lambda}), \\ \sum_{e \in \mathcal{E}(x; \theta)} \tilde{\lambda}^e = 1 \end{array} \right\}. \quad (19)$$

The conditional outcome vector $P(y|x)$ is a point in this space. By definition, $\theta \in \Theta_I$ if and only if $P(y|x) \in \mathcal{P}(\theta)$, the restrictions the statement of the theorem can be rewritten as $\{\theta : P(y|x) \in \mathcal{P}(\theta)\}$. \square

Theorem 5 Given assumptions 1 and 4, the sharp identified set of θ for the full model is,

$$\Theta_I = \{\theta \in \Theta : \forall x \in X, \exists \lambda \in \Lambda(\theta) \text{ s.t. } P(y|x) = \Psi(y|x, \theta, \lambda)\},$$

where,

$$\Lambda(\theta) = \left\{ \lambda : \forall \epsilon, x, \sum_{e \in \mathcal{E}(\epsilon, x, \theta)} \lambda^e(\epsilon, x) = 1 \text{ and } \lambda^e(\epsilon, x) > 0 \right\}.$$

Proof. The proof of this theorem is a trivial extension of the argument from Theorem 4. \square

B Computational Appendix

Computing the Equilibrium Set

In this section I present a method to approximate all equilibrium in the set $\mathcal{E}(\epsilon, x, \theta)$. For simplicity, we suppress covariates in this section and treat μ and δ as constants. Moreover, because we deal with each ϵ draw independently, the notation of this appendix suppresses the dependence of the strategies (χ_1, χ_2) and beliefs (ρ_1, ρ_2) on x and ϵ .

The set of equilibria is equivalent to the set of all solutions to the system of equations,

$$\begin{aligned}\chi_1 &= -(\mu_1(\theta) + \epsilon_1) + \rho_2(\chi_2; \theta)\delta_1(\theta) \\ \chi_2 &= -(\mu_2(\theta) + \epsilon_2) + \rho_1(\chi_1; \theta)\delta_2(\theta)\end{aligned}\tag{20}$$

Where χ_i is a cutoff strategy for entry and ρ_i is agent i 's probability of entry based given he is using the strategy χ_i .

$$\rho_i(\chi; \theta) = \int \mathbf{1}[\nu_i \geq \chi] dF(\nu_i; \theta)$$

We need only search for equilibrium strategies within the set of rationalizable strategies. Because agents beliefs about the probability of entry are bounded between 0 and 1, the set of rationalizable strategies for player i is,

$$\Psi_i = [\min(-(\mu_i(\theta) + \epsilon_i), -(\mu_i(\theta) + \epsilon_i) + \delta_i(\theta)), \max(-(\mu_i(\theta) + \epsilon_i), -(\mu_i(\theta) + \epsilon_i) + \delta_i(\theta))]$$

Equilibrium strategies must be rationalizable. Therefore, we can confine our search for equilibrium cutoffs for player i to Ψ_i . We then search for equilibria using the following algorithm:

1. For a grid of points $\chi_1^p \in \Psi_1$:

(a) Compute player 2's best response given player 1 uses the strategy χ_1^p ,

$$\chi_2^{bp} = -(\mu_2(\theta) + \epsilon_2) + \rho_1(\chi_1^p; \theta)\delta_2(\theta).$$

(b) Compute player 1's best response given player 2 uses the strategy χ_2^{bp}

$$\chi_1^{bp} = -(\mu_1(\theta) + \epsilon_1) + \rho_2(\chi_2^{bp}; \theta)\delta_1(\theta).$$

(c) Compute $z^p = \chi_1^p - \chi_1^{bp}$.

2. Wherever z^p and z^{p+1} are opposite signs, use Newton's method starting (χ_1^b, χ_2^{bp}) to solve (20).

3. If $|z^p| < |z^{p-1}|$ and $|z^p| < |z^{p+1}|$, use Newton's method starting (χ_1^b, χ_2^{bp}) to solve (20).

If (χ_1^b, χ_2^{bp}) is an equilibrium, then $z^p = 0$. The vector of points $\{z^p\}$ is a discretization of a continuous function. The algorithm locates equilibria by searching near the zeros of this function, which is much more efficient than simple multi-starting. Finding all equilibria depends on using a fine enough discretization of the rationalizable set. Clearly, there is a tradeoff between accuracy and computation time. The results of this paper are robust to changing the coarseness of the discretization of the rationalizable set.

Computation of the Sieve Profiled Likelihood Function

This appendix provides details on the optimization problem which is used to calculate the profiled likelihood $L_M(\theta)$. Given a value for λ , the maximand $L_M(\theta, \lambda)$ is calculated using standard numerical simulation techniques. We can write the maximand as,

$$L_M(\theta, \lambda) = \frac{1}{M} \sum_{i=1}^M [\log(\Psi^R(y_i|x_i, \theta, \lambda))]$$

where,

$$\Psi^R(y|x, \theta, \lambda) = \frac{1}{R} \sum_{r=1}^R \sum_{e \in \mathcal{E}(\epsilon_r, x, \theta)} \lambda^e(\epsilon_r, x) \tilde{\Psi}(y_i|\epsilon_r, x, \theta, e) f(\epsilon_r; \theta)$$

Where we have selected R sample points over which to approximate the integral over ϵ . These sample points could be chosen in several different ways, including monte carlo simulation. I have chosen to use Halton sequences to approximate this integral.

Our task is to profile λ out of this function. To accomplish this, we need to find a "most favorable" selection mechanism given θ . Let λ_θ be any element of the set of maximizers of the likelihood for a fixed θ .

$$\lambda_\theta \in \underset{\lambda(\cdot)}{\operatorname{argmax}} L_M(\theta, \lambda(\cdot))$$

It is clear that all λ which are equal on the sample points chosen for the numerical approximation of P^R evaluate to the same likelihood, so λ_θ is not be uniquely defined. However, maximizing λ over the set of sample points will yield an appropriate value for the purpose of approximating $L(\theta)$.

For each sample point I calculate the equilibrium set $\mathcal{E}(\epsilon_r, x, \theta)$. For each equilibrium, I assign each an index e , and a mixing probability $\lambda_{r,x,e}$. We then optimize the following constrained maximization problem over the vector of mixing probabilities,

$$\begin{aligned} \lambda_\theta^R = \underset{\lambda \in [0,1]^{R \times X \times \bar{E}}}{\operatorname{argmax}} & \sum_{i=1}^N \sum_{y \in Y} \mathbf{1}[y_i = y] \log \left(R^{-1} \sum_{r=1}^R \sum_{e \in \mathcal{E}(\epsilon_r, x_i, \theta)} \lambda_{r,x_i,e} \tilde{\Psi}(y|\epsilon_r, x_i, \theta, e) f(\epsilon_r; \theta) \right) \quad (21) \\ \text{s.t. } \forall r, x : & \sum_{e \in \mathcal{E}(\epsilon_r, x, \theta)} \lambda_{r,x,e} = 1 \end{aligned}$$

Although high-dimensional and somewhat daunting in appearance, the optimization problem in (21) is a concave objective with linear constraints, and can be handled by modern nonlinear solvers for R in the hundreds. Furthermore, whenever $\mathcal{E}(\epsilon_r, x, \theta)$ is unique, the selection mechanism is degenerate and there is no need to optimize over the selection mechanism. This leads to a dramatic reduction in the number of unknowns in this problem in many cases. Efficient computation of (21) is important because this problem must be solved for each θ we wish to test during simulated annealing.

We can now evaluate the profiled likelihood statistic for each θ by plugging λ_θ^R back into the full likelihood function.

$$L_M(\theta) = L_M(\theta, \lambda_\theta^R)$$

Weighted Bootstrap Algorithm for Confidence Sets

This appendix describes the implementation of the weighted bootstrap to derive the confidence region for the identified set and the confidence region for the identifiable parameter. Let the likelihood function be defined as,

$$L_M(\theta) = \max_{\lambda \in \Lambda^R} \frac{1}{M} \sum_{i=1}^M \log(\Psi^R(y_i|x_i, \theta, \lambda)),$$

we have described how to compute this function above. The *weighted* likelihood function is defined analogously with the addition of a vector of weights, $w = (w_1, \dots, w_n)$,

$$L_M(\theta, w) = \max_{\lambda \in \Lambda^R} \frac{1}{\sum w_i} \sum_{i=1}^M w_i \log(\Psi^R(y_i|x_i, \theta, \lambda)).$$

Where the weights satisfy $E[w_i^b] = 1$, $V[w_i^b] = 1$ and are independent of the data. Given these assumptions, the weighted likelihood will have the same asymptotic distribution as the “standard” likelihood Chen et al. (2011). Bootstrapping the weighted likelihood amounts to evaluating the function for different sets of weights. The quantiles found by bootstrapping the weighted likelihood will approximate the quantiles of the asymptotic distribution of the likelihood under the null hypothesis.

The likelihood ratio statistic and its weighted bootstrap analogue are defined,

$$LR(\theta) = \max_{\theta'} L_M(\theta') - L_M(\theta),$$

$$LR(\theta, w) = \max_{\theta'} L_M(\theta', w) - L_M(\theta, w).$$

The weighted likelihood and the unweighted likelihood have the same asymptotic distribution, so we can use the quantiles of $\{LR(\theta, w^b)\}_{b=1}^B$ to estimate the cutoff for the confidence region. Ideally, we would use all points in Θ for this procedure, however this is clearly computationally infeasible. Instead we will use simulated annealing to select a large number of points which adequately cover the parameter space near its maximum. Some tuning of the jump distance and the temperature of the simulated annealing algorithm may be needed to ensure adequate coverage.

1. From multiple (around 40) start points, run the simulated annealing algorithm on $LR(\cdot)$ for many (over 10,000) iterations each. Save all points.
2. Define the starting cutoff c_0 and the starting set of points $\mathcal{S}_0 = \{\theta : LR(\theta) \leq c_0\}$. In practice we will use the set of points S_0 which are in \mathcal{S}_0 and have been visited by the simulated annealing. The starting cutoff must be decreasing in n at a slow enough rate a la CHT. An extreme alternative is to let $c_0 = \infty$, which implies $\mathcal{S}_0 = \Theta$ and S is simply all points visited by simulated annealing.
3. For each point in S_0 , compute $\{LR(\cdot, w^b)\}$ note that we only need to solve the model 1 time for each θ and then can compute the likelihood for each weight sample.
4. Iterate the following until $|c_{\ell-1} - c_\ell| < \varepsilon$.

(a) Compute:

$$c_{\ell+1} = \inf\{x : \frac{1}{B} \sum_{b=1}^B \mathbf{1}[\max_{\theta \in S_\ell} (LR(\theta, w^b) - LR(\theta)) \leq x] \geq 1 - \alpha\}.$$

(b) Define:

$$S_{\ell+1} = \{\theta \in S_\ell : LR(\theta) \leq c_{\ell+1}\}.$$

5. Let $\hat{c}_M(1 - \alpha) = c_\ell$, $\hat{\Theta}_I = \{\theta : LR(\theta) \leq \hat{c}_M(1 - \alpha)\}$. Use S_ℓ to construct confidence intervals for statistics of interest. Such confidence intervals are valid assuming the number of simulated annealing iterations goes to ∞ with n .

To find the confidence set for the identifiable parameter we individually test the hypothesis that $\theta \in \Theta_I$ for each point. The collection of all points that are not rejected is denoted $\bar{\Theta}$, the confidence set for the identifiable parameter. It is easy to see that any point in $\bar{\Theta}$ must also be in $\hat{\Theta}_I$ —if we cannot reject that θ is the true parameter, we clearly cannot exclude it from our coverage region for the identified set. Therefore, we can restrict the hypothesis test to all points in $\hat{\Theta}_I$. The test is conducted for a given θ by computing,

$$\hat{c}_n(\theta, 1 - \alpha) = \inf\{x : \frac{1}{B} \sum_{b=1}^B \mathbf{1}[LR(\theta, w^b) - LR(\theta) \leq x] \geq 1 - \alpha\}.$$

The hypothesis is rejected if $LR(\theta) > \hat{c}_n(\theta, 1 - \alpha)$. So the confidence set for the identifiable parameter is,

$$\bar{\Theta} = \{\theta \in \hat{\Theta}_{CR} : LR(\theta) \leq \hat{c}_n(\theta, 1 - \alpha)\}.$$

Again, we use the set of points visited by simulated annealing which pass the above condition to construct confidence intervals for the statistics of interest.

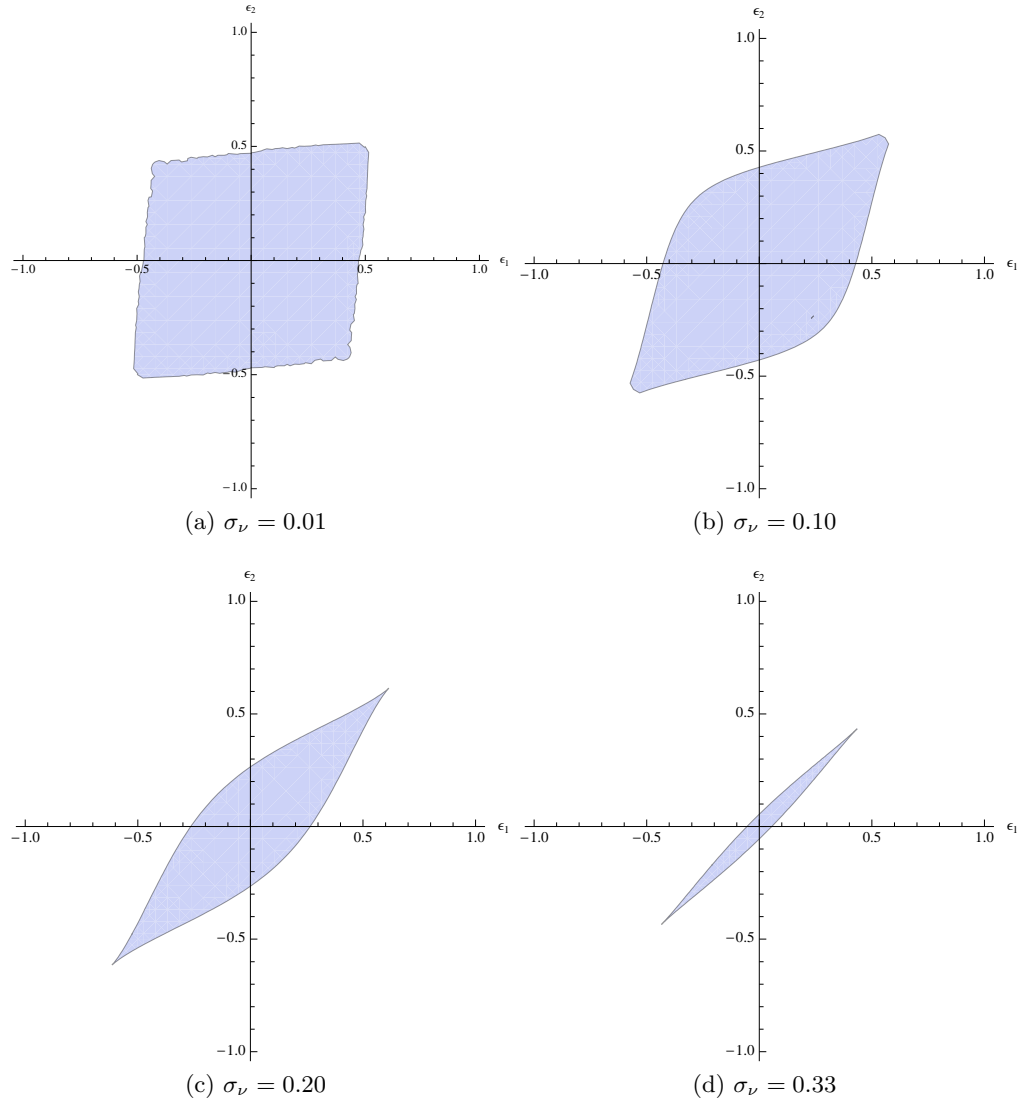


Figure 1: Multiplicity in a 2-player game correspondence $\mathcal{E}(\cdot, \theta)$ varying the degree of incomplete information by panel. There are multiple equilibria in the shaded region, and one equilibrium in the unshaded region. Axes correspond to (ϵ_1, ϵ_2) . For both players $\theta_\mu = 0.5$ and $\theta_\delta = -1$. The region of multiplicity for the limiting complete information game is the box $[-0.5, 0.5] \times [-0.5, 0.5]$ (cf. Bresnahan and Reiss, 1991a, Figure 1).

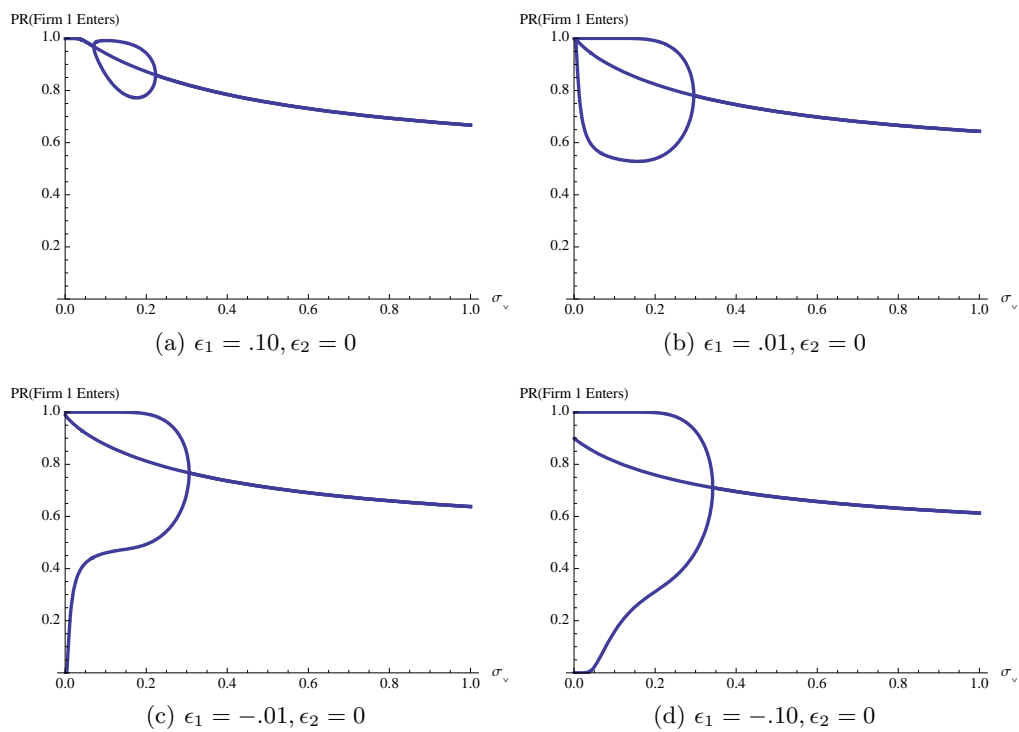


Figure 2: Equilibrium entry probability of player 1 in a symmetric two-player entry game as the level of uncertainty changes. For both players, $\mu_i = \delta_i = 1$.

	Quantiles			Mean	Std Dev
	0.25	0.50	0.75		
Population	2,610	4,327	6,700	4,981.04	3,109.26
Mean Household Income	38,013	43,135	50,139	45,526.28	12,235.60
Distance to Supercenter					
1998	17.15	30.67	77.81	76.51	113.22
2000	14.64	24.12	49.40	53.74	80.41
2002	13.28	20.79	33.67	33.55	39.12
Market in South				0.35	0.47
Market in West				0.14	0.34

Table 1: Summary statistics of market characteristics. Mean distance to a supercenter excludes outliers more than 500 miles from a supercenter. These are accounted for in the quantile calculation. Total number of markets is 4,803.

Active Firms	Supercenter		Overall
	Yes	No	
0	22.41	23.09	22.82
1	65.37	60.50	62.46
2	12.22	16.41	14.72
N	1,932	2,871	4,803

Table 2: Distribution (of the number) of active firms in 2002 by whether the market was within 20 miles of a supercenter in 2000 (percent).

	Percent Near Supercenter	Active Firms		
		0	1	2
Pop 0-3k	25.91			
Supercenter		48.16	49.47	2.07
No supercenter		43.66	53.62	2.72
Pop 3-6k	42.47			
Supercenter		22.99	70.78	6.23
No supercenter		13.90	71.62	14.48
Pop 6k+	51.73			
Supercenter		9.02	67.78	23.20
No supercenter		4.97	54.97	40.06

Table 3: Distribution (of the number) of active firms in 2002 by market size and whether or not a supercenter was located within 20 miles in 2000 (percent).

	I	II	III	IV	V
Supercenter within 20 mi	-0.0834 (0.0493)	-0.0311 (0.0565)	-0.0897 (0.0496)	-0.0429 (0.0569)	-0.0862 (0.0498)
Supercenter entry 1998-2000		-0.1454 (0.0882)		-0.1392 (0.0883)	
Log Population	0.3358 (0.0352)	0.3308 (0.0371)			
Log Household Income		-0.0405 (0.1037)		-0.0694 (0.1032)	
South Region		0.0523 (0.0544)		0.0415 (0.0546)	
West Region		0.1207 (0.0687)		0.0828 (0.0692)	
Pop over 3k			-0.4248 (0.0595)	-0.4222 (0.0613)	-0.3680 (0.0979)
Pop over 6k			0.6992 (0.0657)	0.6955 (0.0695)	0.5648 (0.0904)
Competitor Active in 1998	-0.5076 (0.0509)	-0.5011 (0.0512)	-0.5047 (0.0514)	-0.4999 (0.0517)	-0.7406 (0.1051)
(Pop > 3k)&(Comp. in 98)					-0.0735 (0.1230)
(Pop > 6k)&(Comp in 98)					0.3653 (0.1406)
Log-Likelihood	-1758.17	-1755.10	-1748.36	-1746.05	-1744.46
N	5217	5217	5217	5217	5217

Table 4: Probit regressions on firm entry for five year period from 1998 to 2002. These results do not control for endogeneity of decisions between small grocery stores.

	I	II	III	IV	V
Supercenter within 20 mi	0.1463 (0.0508)	0.1501 (0.0575)	0.1500 (0.0511)	0.1558 (0.0578)	0.1567 (0.0512)
Supercenter entry 1998-2000		-0.1314 (0.0939)		-0.1298 (0.0939)	
Log Population	-0.3043 (0.0387)	-0.3136 (0.0395)			
Log Household Income		0.0932 (0.1136)		0.0894 (0.1146)	
South Region		0.0160 (0.0556)		0.0160 (0.0557)	
West Region		-0.1044 (0.0778)		-0.0923 (0.0777)	
Pop over 3k			0.1941 (0.0586)	0.2027 (0.0597)	0.1989 (0.0786)
Pop over 6k			-0.5199 (0.0711)	-0.5351 (0.0730)	-0.5779 (0.0841)
In Duopoly in 1998	0.2377 (0.0547)	0.2394 (0.0550)	0.2348 (0.0557)	0.2366 (0.0561)	-0.0915 (0.1663)
(Pop > 3k)&(Doup. in 98)					0.0214 (0.1176)
(Pop > 6k)&(Doup. in 98)					0.3609 (0.1852)
Log-Likelihood	-1642.94	-1640.53	-1646.84	-1644.69	-1644.53
N	4389	4389	4389	4389	4389

Table 5: Probit regressions on firm exit for five year period between 1998 and 2002. These results do not control for endogeneity of decisions between small grocery stores.

	I	II
Monopoly Profits, $\mu(\cdot)$		
Constant	0.219 (0.033)	0.196 (0.035)
$\mathbf{1}[Pop > 3k]$	0.827 (0.043)	0.845 (0.049)
$\mathbf{1}[Pop > 6k]$	0.704 (0.043)	0.574 (0.060)
$\mathbf{1}[Supercenter < 20mi]$	-0.336 (0.036)	-0.255 (0.045)
Competition Effect, $\delta(\cdot)$		
Constant	-2.101 (0.033)	-2.054 (0.073)
$\mathbf{1}[Pop > 3k]$		-0.059 (0.088)
$\mathbf{1}[Pop > 6k]$		0.227 (0.074)
$\mathbf{1}[Supercenter < 20mi]$		-0.187 (0.063)
Log-Likelihood	-3846.3	-3838.3
N	4803	4803

Table 6: Results from a Bresnahan-Reiss complete information model with symmetric firms.

Parameter	Incomplete Info.		Complete Info.	Full Model	
	Identified Set	Identifiable Parameter		Identified Set	Identifiable Parameter
Monopoly Profits, $\mu(\cdot)$					
Constant	-1.222 (0.060)	-1.233 (0.040)		[-1.764 -0.921]	[-1.578 -0.974]
$\mathbf{1}[Pop > 3k]$	0.293 (0.062)	0.215 (0.042)		[-0.067 0.652]	[0.094 0.584]
$\mathbf{1}[Pop > 6k]$	0.544 (0.072)	0.538 (0.069)		[0.248 1.347]	[0.356 1.281]
$\mathbf{1}[Supercenter < 20mi]$	-0.073 (0.055)	-0.052 (0.044)		[-0.430 0.240]	[-0.269 0.116]
Entry Cost	2.158 (0.040)	2.221 (0.029)		[1.630 2.527]	[1.630 2.419]
Competition Effect, $\delta(\cdot)$					
Constant	-0.851 (0.140)	-0.573 (0.057)		[-1.663 0.056]	[-1.599 -0.150]
$\mathbf{1}[Pop > 3k]$	0.181 (0.133)	0.092 (0.061)		[-0.804 0.902]	[-0.420 0.767]
$\mathbf{1}[Pop > 6k]$	0.368 (0.136)	0.108 (0.080)		[-0.680 0.699]	[-0.421 0.530]
$\mathbf{1}[Supercenter < 20mi]$	-0.148 (0.083)	-0.078 (0.056)		[-0.610 0.401]	[-0.478 0.279]
Public v. Private Info, γ	0.000	1.000		[0.347 1.000]	[0.446 1.000]
Correlation of Public Info, ρ	n.a.	-0.068 (0.030)		[-0.972 0.796]	[-0.933 0.780]
Log-Likelihood	-3400.0	-3296.7		-3338.3 to -3289.3	
In CR of Identified Set, $\theta \in \hat{\Theta}$	No	Yes			
In CR of Identifiable Parameter, $H_0 : \theta = \theta^*$	Rejected	Not Rejected			

Table 7: Results from the structural model. The incomplete information model has a unique equilibrium at the objective optimum, so results are reported assuming uniqueness. For the complete information model, I assume that a pure equilibrium is chosen with probability .5 to resolve multiplicity with this assumption the model is point identified. For the full, this table reports projections of the 95 percent joint confidence region onto the parameter axes. The true confidence region is a subset of the box reported in the table. Their are 4,803 market observations

Effect of Supercenter Entry on Firm Value (Percent)		
	Complete Info.	Full Model
Pop 0-3k		
Monopoly	[-10.7 2.5]	[-18.7 9.3]
Duopoly	[-10.6 -2.4]	[-18.8 -0.1]
Pop 3k-6k		
Monopoly	[-9.6 2.3]	[-16.0 7.6]
Duopoly	[-11.0 -3.1]	[-22.1 0.7]
Pop 6k+		
Monopoly	[-8.2 1.6]	[-12.7 5.9]
Duopoly	[-10.8 -3.2]	[-24.6 1.2]

Table 8: Effect on firm value of grocery store if Supercenter enters within 20 miles (percent), 95 percent confidence intervals for the complete information model and the full model.

Effect of Local Entry on Firm Value (Percent)		
	Complete Info.	Full Model
Pop 0-3k		
Supercenter	[-28.1 -22.4]	[-59.9 -8.3]
No Supercenter	[-26.2 -18.7]	[-50.8 -10.7]
Pop 3k-6k		
Supercenter	[-27.3 -14.5]	[-47.8 -10.1]
No Supercenter	[-24.9 -21.1]	[-58.1 -3.6]
Pop 6k+		
Supercenter	[-24.5 -13.8]	[-55.9 -9.2]
No Supercenter	[-23.0 -10.2]	[-46.0 -6.3]

Table 9: Effect on a monopolist grocery store’s firm value when another grocery store enters its market (percent); 95 percent confidence intervals.

Difference in Store Value Following Supercenter versus Local Entry		
	Complete Info.	Full Model
Pop 0-3k	[12.6 25.6]	[-5.0 43.4]
Pop 3k-6k	[10.4 20.1]	[-2.1 45.2]
Pop 6k+	[6.4 19.9]	[-1.0 39.1]

Table 10: Difference between the value of a store following supercenter entry versus local grocery store entry as a percentage of the monopolist store value. A positive number implies that a monopolist would prefer supercenter entry to local-grocery-store entry, a negative number implies the opposite; 95 percent confidence intervals.

	Number of Grocery Stores		
	0	1	2
Pop 0-3k			
Supercenter	[30.3 62.1]	[37.0 68.8]	[0.2 4.6]
No Supercenter	[26.3 51.7]	[45.4 70.9]	[0.5 8.2]
Pop 3k-6k			
Supercenter	[14.1 34.9]	[59.2 79.0]	[2.2 12.2]
No Supercenter	[10.0 28.0]	[63.0 81.2]	[5.7 21.2]
Pop 6k+			
Supercenter	[2.1 14.1]	[59.6 82.5]	[11.4 34.0]
No Supercenter	[1.2 8.0]	[40.8 73.6]	[22.8 57.5]

Table 11: Stationary distribution of the number of firms in a market by market type (percent); 95 percent confidence intervals.

	Expected Number of Grocery Stores	Supercenter Effect (Percent)
Pop 0-3k		
Supercenter	[0.39 0.71]	
No Supercenter	[0.50 0.81]	[-36.9 15.4]
Pop 3k-6k		
Supercenter	[0.69 0.96]	
No Supercenter	[0.81 1.08]	[-25.1 0.6]
Pop 6k+		
Supercenter	[0.99 1.28]	
No Supercenter	[1.17 1.56]	[-27.9 1.0]

Table 12: Expected number of stores by market size and supercenter presence and the effect of adding a supercenter to a market on the expected number of stores; 95 percent confidence intervals.