Empirical Models of Entry and Market Structure \(^1\)

Steven Berry  
Yale Univ. and NBER  
steven.berry@yale.edu

Peter Reiss  
Stanford Univ. and NBER  
preiss@optimum.stanford.edu

February 5, 2006

\(^1\)Draft Chapter for Volume III of the *Handbook of Industrial Organization*. Please do not circulate or cite without the authors’ permission.
Abstract

This chapter considers the estimation of equilibrium models of endogenous market structure. We pay particular attention to simple “entry” models estimated on cross-sectional data. We review estimation methods and also discuss some illustrative empirical examples. The review begins with simple symmetric models and then discusses various approaches to dealing with the multiple equilibria that often arise in models with heterogeneous firms. We also briefly discuss some recent work on models with incomplete information and work on “set identified” models of entry.
1 Introduction

Industrial organization (IO) economists have devoted substantial energy to understanding market structure and the role that it plays in determining the extent of market competition. In particular, IO economists have explored how the number and organization of firms in a market, firms’ sizes, potential competitors, and the extent of firms’ product lines affect competition and firm profits. This research has shaped the thinking of antitrust, regulatory and trade authorities who oversee market structure and competition policies. For example, antitrust authorities regularly pose and answer the question: How many firms does it take to sustain competition in this market? Others ask: Can strategic investments in R&D, advertising and capacity deter entry and reduce competition? Firms themselves are interested in knowing how many firms can ‘fit’ in a market.

Not too surprisingly, economists’ thinking about the relationship between market structure and competition has evolved considerably. Theoretical and empirical work in the 1950s, 1960s and early 1970s examined how variables such as firm profits, advertising, R&D, and prices differed between concentrated and unconcentrated markets. Much of this work implicitly or explicitly assumed market structure was exogenous. Early efforts at explaining why some markets were concentrated, and others not, relied on price theoretic arguments that emphasized technological differences and product market differences (e.g., Bain (1956)). In the 1970s and 1980s, IO models focused on understanding how strategic behavior might influence market structure. Much of this work treated market structure as the outcome of a two-stage game. In a first stage, potential entrants would decide whether to operate; in a second stage, entering firms would compete along various dimensions. Although these two-stage models underscored the importance of competitive assumptions, the predictions of these models sometimes were very sensitive to specific modeling assumptions.

Also during the 1970s and 1980s the increased availability of census and other micro-panel data allowed empiricists to document rich and diverse patterns of firm turnover and changes in industry structures. For example, T. Dunne and Samuelson (1989) study revealed considerable heterogeneity in firm survival by type of entry and significant positive cross-industry correlations in entry and exit rates.

The richness of these theoretical models and firm-level data sets have presented

\footnote{The term market structure broadly refers to the number of firms in a market, their sizes and the products they offer.}
empiricists with a unique opportunity to develop empirical models capable of evaluating competing theories of industrial concentration. With this opportunity, however, comes the practical challenge of linking the existing theories to the available data. The process of linking analytical market structure models to data involves three critical steps:

1. determining what alternative theories say about observable price, quantity, and entry and exit data;
2. specifying how unobservables in the theory impact the distribution of the observed data;
3. resolving any ambiguities in the models, including the possibility of no equilibria or multiple equilibria.

This chapter has two broad goals. The first is to illustrate how empirical IO economists can use game-theoretic models to build structural econometric models of entry, exit and market concentration. Part of our discussion seeks to describe what interesting economic questions empirical models can address with readily available data. Our discussion also seeks to examine how empirical predictions about market structure depend on observable and unobservable economic variables, including:

- the size and sunkness of set-up costs;
- the sensitivity of firm profits to the entry and exit of competitors;
- the extent of product substitutability and product lines;
- potential entrants’ expectations about post-entry competition;
- the (unseen) supply of potential entrants; and
- the endogeneity of fixed and sunk costs.

We begin by outlining an econometric framework suitable for analyzing cross-section variables such as the number and sizes of firms in a market. This framework treats the number and identity of firms as endogenous outcomes of a two-stage oligopoly game. In the first stage, firms decide whether to operate and perhaps some product characteristics, such as quality; in the second stage, the entering firms compete (the
nature of this competition may be made specific or may be left as a kind of “reduced form”). This simple framework allows us to consider various economic questions about the nature of competition and sources of firm profitability.

A second goal of this chapter is to provide some sense of where the IO literature is in empirically modeling the economic significance of market structure. We should note that we do not attempt to survey papers that summarize turnover patterns in different industries or sectors. Several excellent surveys of these literatures already exist, including Geroski (1995) and Caves (1998). Instead, we review and interpret existing structural econometric models. Most of the papers we discuss estimate the parameters of static two-stage oligopoly models using cross-section data covering different geographic markets. In this sense, these models are more about modeling long-run equilibria. Although there are a few studies that analyze the structure of the same market through time, these typically do not model the endogenous timing of entry and exit decisions. Later in this chapter, we discuss special issues that time-series data and dynamic models pose for modeling market structure and changes in market structure. As the chapter by Ariel Pakes suggests, however, explicit dynamic models in which firms engage in strategic behavior often raise some very challenging econometric and computational issues. Thus, while these models are more theoretically appealing, they are not easily applied to commonly available data. To date, there have been relatively few attempts at estimating dynamic models.

In this chapter, we emphasize models, economic questions and empirical methodology. We also provide several short discussions of actual empirical applications. While we provide some intuitive discussion of the sources of variation in the data that help us to identify the parameters of the model, we leave most a formal treatment of non-parametric identification to the work described in Berry and Tamer (2006).²

1.1 Why Structural Models of Market Structure?

The primary reason we focus on structural models of market structure, entry or exit is that they permit us to estimate unobservable economic quantities that we could not recover using descriptive models. For instance, to assess why a particular market is concentrated, IO economists must distinguish between fixed and variable cost explanations. Unfortunately, IO economists rarely, if ever, have the accounting or other data necessary to construct accurate measures of firms’ fixed or variable costs.

²See also the semi-parametric approach of Jia (2003) and Bajari, Hong, and Ryan (2004).
This means that IO economists must use other information, such as prices, quantities and the number of firms in a market, to recover information about demand, variable costs and fixed costs. As we shall see, to use information on prices, quantities and market concentration to recover information on firms profits, researchers must often make stark modeling assumptions. In some cases, small changes in assumptions, such as the timing of firms’ moves or the game’s solution concept, can have a dramatic effect on inferences. In the framework that follows, we illustrate some of these sensitivities by comparing models that use plausible alternative economic assumptions. Our intent in doing so is to provide some feel for which economic assumptions are crucial to inferences about market structure and unobservables such as fixed costs, variable costs, demand, and strategic interactions.

When exploring alternative modeling strategies we also hope to illustrate how practical considerations, such data limitations, can constrain what the researcher can identify. For example, ideally a researcher would know who is a potential entrant, firms’ expectations about competitors, entrants’ profits, etc. In practice, IO researchers rarely have this information. For instance, they may only observe who entered and not who potentially could have. In general, the less the IO economist knows about potential entrants, firms’ expectations, entrants’ profits, etc, the less they will be able to infer from data on market structure, and the more the researcher will have to rely on untestable modeling assumptions.

Many of these general points are not new to us. Previously Bresnahan (1989) reviewed ways in which IO researchers have used price and quantity data to draw inferences about firms’ unobserved demands and costs, and competition among firms. Our discussion complements his and other surveys of the market power literature in that most static entry models are about recovering the same economic parameters. There are some key differences however. First, the market power literature usually treats market structure as exogenous. Second, the market power literature does not try to develop estimates of firms’ fixed, sunk, entry or exit costs. Third, market structure studies occasionally do not have the price and quantity information necessary to conduct an analysis of market power.

From a methodological point of view, market structure models improve on market power models in that they endogenize the number of firms in a market. They do this by simultaneously modeling potential entrants (discrete) decisions to enter or not enter a market. These models rely on the insight that producing firms expected non-negative economic profits, conditional on the actions of competitors, including those who did not enter. This connection is analogous to the revealed preference
arguments that form the basis for discrete choice models of consumer behavior. As in the consumer choice literature, firms' discrete entry decisions are interpreted as revealing something about an underlying latent profit or firm objective function. By observing how firms' decisions change, as their choice sets and market conditions change, IO economists can gain insight into the underlying determinants of firm profitability, including perhaps the role of fixed (and/or sunk) costs, the importance of firm heterogeneity, and the nature of competition itself.

If firms made entry decisions in isolation it would be a relatively simple matter to adapt existing discrete-choice models of consumer choice to firm entry decisions. In concentrated markets, however, firms entry decisions are interdependent – both within and sometimes across product markets. These interdependencies considerably complicate the formulation and estimation of market structure models. In particular, the interdependence of discrete entry decisions can pose thorny identification and estimation problems. These problems generally cannot be assumed away, without altering the realism of the model of firm decision making.

For example, simultaneous discrete-choice models are known to have “coherency” problems (e.g., Heckman (1978)) that can only be fixed by strong statistical assumptions. The industrial organization literature that we describe in this chapter has adopted the alternative approach of asking what combinations of economic and statistical assumptions can ameliorate these problems. Similar approaches arise and are being tried in household labor supply models. (See for example Bjorn and Vuong (1984), Kooreman (1994) and Bourguignon and Chiappori (1992).) While we emphasize the sources of variation in data that can help to identify entry models, we leave a formal treatment of non-parametric identification to other work, including Berry and Tamer (2006).

In what follows, we start with simple illustrative models and build toward more complex ones.

2 Entry Games with Homogeneous Firms

This section outlines how IO economists have used the number of firms in a market to recover information about market demand, and firms’ costs. It does so under the untenable assumption that all potential entrants are the same. The advantage of this assumption is that it allows us to isolate general issues that are more difficult to appreciate in complicated models. The section that follows relaxes the untenable
2.1 A Simple Homogeneous Firm Model

Our goal is to develop an empirical model of $N$, the number of homogeneous firms that choose to produce a homogeneous good. To do this, we develop a two-period oligopoly model in which $M$ potential entrants first decide whether to enter and then how much to produce, $q$. In developing the empirical model, we limit ourselves to the all too common situation where the empiricist observes $N$ but not firm output $q$.

The empirical question we seek to address with this model is: What can we learn about economic primitives, such as demand, cost and competitive behavior from observations on the number of firms $N_1, ..., N_T$ that entered $T$ different markets. To do this, we need to relate $N_i$ to firm profits in market $i$. Given $N_i$, each entrant in market $i$ earns

$$
\pi(N_i) = V(N_i, x_i, \theta) - F_i.
$$

Here, $V(\cdot)$ represents a firm’s variable profits and $F$ is a fixed cost. Under our homogeneity assumption, all firms in market $i$ have the same variable profit function and fixed cost $F_i$. The vector $x_i$ contains market $i$ demand and cost variables that affect variable profits. The vector $\theta$ contains the demand, cost and competition parameters that we seek to estimate. To relate this profit function to data on the number of firms, we assume that in addition to observing $N_i$, we observe $x_i$ but not $\theta$ or fixed costs $F_i$. While in principle $x_i$ could include endogenous variables such as prices or quantities, we simplify matters for now by assuming that the only endogenous variable is the number of firms that entered $N_i$.

Before estimation can proceed, the researcher must specify how variable profits depend on $N$ and what is observable to the firms and researcher. These two decisions typically cannot be made independently, as we shall see in later subsections. The purpose of explicitly introducing variables that the econometrician does not observe is to rationalize why there is not an exact relation between $x_i$ and $N_i$. To make matters simple, we assume that firms have complete information about each other’s profits, but that the researcher does not observe firms’ fixed costs. Further, for simplicity we assume that the fixed costs the econometrician does not observe are independently distributed across markets according to the distribution $\Phi(F|x, \omega)$. As the sole source of unobservables, the distribution $\Phi(F|x, \omega)$ describes not only the distribution of $F_i$, but firm profits $\pi(N_i)$ as well.
Once the profit function is in place, the researcher’s next task is to link firms’ equilibrium entry decisions to $N$. Because firms are symmetric, have perfect information and their profits are a nonincreasing function of $N^*$, we only require two inequalities to do this. For the $N^*$ firms that entered

$$V(N^*, x, \theta) - F \geq 0,$$  \hspace{1cm} (2)

and for any of the other potential entrants

$$V(N^* + 1, x, \theta) - F < 0.$$  \hspace{1cm} (3)

When combined, these two equilibrium conditions place an upper and lower bound on the unobserved profits

$$V(N^*, x, \theta) \geq F > V(N^* + 1, x, \theta).$$  \hspace{1cm} (4)

These bounds provide a basis for estimating of the variable profit and fixed cost parameters $\theta$ and $\omega$ from information on $x_i$ and $N_i$. For example, we use the probability of observing $N^*$ firms:

$$\text{Prob}(V(N^*, x) \geq F | x) - \text{Prob}(V(N^* + 1, x) > F | x) = \Phi(V(N^*, x, \theta) | x) - \Phi(V(N^* + 1, x | x).$$  \hspace{1cm} (5)

to construct a likelihood function for $N^*$. Under the independent and identical sampling assumptions we adopted, this likelihood has a practical “ordered” dependent variable form:

$$\mathcal{L}(\theta, \omega | \{x, N^*\}) = \sum_i \ln (\Phi(V(N_i^*, x_i)) - \Phi(V(N_i^* + 1, x_i))$$  \hspace{1cm} (6)

where the sum is over the cross-section or time-series of independent markets in the sample. It is essential to note that besides maintaining that firms’ unobserved profits are statistically independent across markets, this likelihood function presumes that firms’ profits are economically independent across markets. These independence assumptions are much more likely to be realistic if we are modeling a cross section of different firms in different markets, that the same firms over time or in different markets.

The simplicity of this likelihood function, and its direct connection to theory, is extremely useful despite the stringent economic assumptions underlying it. Most important, we learn that if we only have discrete data on the number of firms in different
markets, we will be forced to impose distributional assumptions on unobserved profits in order to estimate $\theta$ and $\omega$. For example, if we assume that unobserved fixed costs have an independent and identically distributed (i.i.d.) normal (logit) distribution, then (6) is an ordered probit (logit) likelihood function. But this begs the questions: How did we know profits had a normal (logit) distribution? And, How did we know that fixed costs were i.i.d. across markets? In general, economics provides little guidance about the distribution of fixed costs. Thus, the message is that absent some statistical (or economic) structure we will be unable to recover much from observations on $N^*$ alone.

Given the potential arbitrariness of the assumptions about fixed costs, it seems imperative that researchers explore the sensitivity of estimates to alternative distributional assumptions. Toward this end, recent work on semiparametric estimators by, among others, Klein and Sherman (2002) and Lewbel (2002), may prove useful in estimating models that have more flexible distributions of unobserved profits. These semiparametric methods, however, typically maintain that $V(\cdot)$ is linear in its parameters and they require large amounts of data in order to recover the distribution of unobserved fixed costs.

To summarize our developments to this point, we have developed an econometric model of the number of firms in a market from two equilibrium conditions on homogeneous firms’ unobserved profits. One condition is based on the fact that $N^*$ chose to enter. The other on the fact that $M - N^*$ chose not to enter. The resulting econometric threshold models bear a close relation to conventional ordered dependent variable models, and thus provide a useful reference point for modeling data on the number of firms. Our discussion also emphasized that to draw information from the number of firms alone, researchers will have to make strong economic and statistical assumptions. In what follows, we show how many of these assumptions can be relaxed, but not without some cost to the simplicity of the econometric model. The next section discusses, following Bresnahan and Reiss (1991b), how one might make inferences about competition in the context of particular models for $V$. We then discuss how to combine information post-entry outcomes with information on entry.

2.2 Relating $V$ to the Strength of Competition.

We now take up the question of how to specify the variable profit $V(N, x, \theta)$ and the fixed cost function $F(x, \omega)$. 
There are two main approaches to specifying how \( V(\cdot) \) depends on \( N \) and \( x \). The first is to pick a parameterization of \( V(\cdot) \) that makes estimation simple and yet obeys the restriction that \( V(\cdot, x) \) is non-increasing in \( N \). For example, the researcher might assume that \( x \) and \( 1/N \) enter \( V \) linearly with constant coefficients, and that the coefficient on \( 1/N \) is constrained to be positive. The advantage of this descriptive approach is that it yields a conventional probit model. The disadvantage is that it is unclear what economic parameters the \( \theta \) represent.

A second approach is to derive \( V(\cdot) \) directly from specific assumptions about the functional forms of demand and costs, and assumptions about the post-entry game. This approach has the advantage of making it clear what economic assumptions motivate the researcher’s choice of \( V(\cdot) \) and what the parameters \( \theta \) represent. A potential disadvantage of this approach is that, even with strong functional form restrictions, the profit specifications can quickly become econometrically challenging.

To see some of the issues involved, consider the market for a homogeneous good with \( M \) potential entrants. Suppose each potential entrant \( j \) has the variable cost function \( C_j(q_j) \) and that demand in market \( i \) has the form

\[
Q_i = S_i \cdot q(P_i),
\]

where \( S \) is market size (an exogenous “\( x \)”), \( q \) is per-capita demand and \( P \) is price. In a standard Cournot model, each entering firm maximizes profits by choosing output so that in market \( i \)

\[
P_i = \frac{\eta_i}{\eta_i - s_j} \cdot MC_j \quad \text{for } j = 1, ..., N \leq M
\]

where \( MC_j \) is entrant \( j \)'s marginal cost of production, \( s_j \) is firm \( j \)'s market share (equal to \( 1/N \) in the symmetric case) and and \( \eta_i \) equals minus the market \( i \) elasticity of demand.\(^3\)

\[^3\] We could extend this to incorporate different possible equilibrium notions in the “usual” way by writing the pricing equation as

\[
P_i = \frac{\eta_i}{\eta_i - \omega_j s_j} \cdot MC_j \quad \text{for } j = 1, ..., N \leq M
\]

where the variable \( \omega_j \) is said to describe firm \( j \)'s “beliefs” about the post-entry game. The usual values are \( \omega_j = 0 \) (competition) and \( \omega_j = 1 \) (Cournot). Current practice is to not think of \( \omega_j \) as an arbitrary conjectural parameter. One could also embed monopoly outcomes within this framework provided we resolve how the cartel distributes production.
As equation (8) stands, it is hard to see how prices (and hence firm-level profits) vary with the number of firms in a market. To explore the effect of \(N\) on prices, it is useful to aggregate the markup equations across firms to give the price equation:

\[
P = \frac{N\eta}{N\eta - 1} \overline{MC}.
\]

(10)

where \(\overline{MC}\) is the average of the \(N\) entrants’ marginal cost functions. This equation shows that industry price depends not just on the number of firms \(N\) that enter the market, but also on the average of the firms’ marginal costs. Alternatively, if interest centers on the size distribution of entrants, we can aggregate (8) using market share weights to obtain

\[
P = \frac{\eta}{\eta - H}\overline{MC}_{aw},
\]

(11)

which links the industry Herfindahl index \(H\) to prices and a market-share weighted average of firms’ marginal costs.

It is tempting to do comparative statics on equations (10) and (11) to learn how entry affects price and variable profits. For example, if we specialize the model to the case where firms: have constant marginal costs, face a constant elasticity of demand and are Cournot competitors, then we obtain the usual “competitive” pattern where prices and variable profits are convex to the origin and asymptote to marginal cost. At this level of generality, however, it is unclear precisely how \(V(\cdot)\) depends on the number of firms. To learn more, we have to impose more structure.

Suppose, for example, that we assumed demand was linear in industry output:

\[
Q = S (\alpha - \beta P).
\]

(12)

Additionally, suppose costs are quadratic in output and the same for all firms

\[
F + C(q) = F + cq - dq^2.
\]

(13)

With these demand and cost assumptions, and assuming firms are Cournot-Nash competitors, we can derive an expression for equilibrium price

\[
P = a - N^* \frac{(a - c)}{(N^* + 1 + 2Sd/b)}.
\]

(14)

Substituting this expression back into demand, we obtain an expression for firm profits:

\[
\pi_i(N^*_i, S_i) = V(N^*_i, S_i, \theta) - F_i = \theta_1^2 S_i \frac{(1 + \theta_2 S^2_i)}{(N^*_i + 1 + 2\theta_2 S_i)^2} - F_i,
\]

(15)
where \( \theta_1 = (a - c)/\sqrt{b} \) and \( \theta_2 = d/b \). This expression can now be inserted into the inequalities \( 4 \) to construct an ordered dependent variable for the number of firms. For example, setting \( d = 0 \) we can transform equation \( 4 \) into

\[
\ln(N_i^* + 2) > \frac{1}{2} \left( \ln(\theta_3 \theta_1 S_i) - \ln F_i \right) \geq \ln(N_i^* + 1). \quad (16)
\]

The identification of the demand and cost parameters (up to scale of unobserved profits or fixed costs) then rests on what additional assumptions we make about whether the demand and cost parameters in \( \theta \) vary across markets and what we assume about the observed distribution of fixed costs.

As should be clear from this discussion, changes in the demand and cost specifications will change the form of the bounds. For example, if we had assumed a unit constant-elasticity demand specification \( P = \theta_1 \frac{S_i}{Q} \) and \( d = 0 \), then we would obtain

\[
V(N_i, S_i) = \theta_1 S_i \frac{1}{N_i^2},
\]

which is similar to that in Berry (1992), and bounds linear in the natural logarithm of \( N \)

\[
\ln(N^* + 1) > \frac{1}{2} \left( \ln(\theta_1 S) - \ln F \right) \geq \ln(N^*). \quad (17)
\]

In this case, knowledge of \( F \) and \( N \) would identify the demand curve. This, however, is not the case in our previous example \( 16 \). More generally, absent knowledge of \( F \), knowledge of \( N \) alone will be insufficient to identify separate demand and cost parameters.

These examples make three important points. First, absent specific functional form assumptions for demand and costs, the researcher will not in general know how unobserved firm profits depend on the number of homogeneous firms in a market. Second, specific functional form assumptions for demand, costs and the distribution of fixed costs will be needed to uncover the structure of \( V(N^*, x, \theta) \). In general, the identification of demand and cost parameters in \( \theta \) separately from \( F \) will have to be done on a case-by-case basis. Finally, apart from its dependence on the specification of demand and costs, the structure of \( V(N^*, x, \theta) \) will depend on the nature of firm interactions. For example, the analysis above assumed firms were Cournot-Nash competitors. Suppose instead we had assumed firms were Bertrand competitors. With a homogeneous product, constant marginal costs and symmetric competitors, price would fall to marginal cost for \( N \geq 2 \). Variable profits would then independent of \( N \).
With symmetric colluders and constant marginal costs, price would be independent of \( N \), and \( V(N) \) would be proportional to \( 1/N \).

Our emphasis on deriving how \( V(\cdot) \) depends on the equilibrium number of firms is only part of the story. Ultimately, \( N \) is endogenous and this then raises an “identification” issue. To see the identification issue, imagine that we do not have sample variation in the exogenous variables \( x \) (which in our examples is \( S \), the size of the market). Without variation in \( x \), we will have no variation in \( N^* \), meaning that we can at best place bounds on \( \theta \) and fixed costs. Thus, \( x \) plays a critical role in identification by shifting variable profits independently of fixed costs. In our example, it would thus be important to have meaningful variation in the size of the market \( S \). Intuitively, such variation would reveal how large unobserved fixed costs are relative to the overall size of the market. In turn, the rate at which \( N \) changes with market size allows us to infer how quickly \( V \) falls in \( N \).

We should emphasize that so far our discussion and example have relied heavily on the assumption that firms are identical. Abandoning this assumption, as we do later, can considerably complicate the relationship between \( V, x \) and \( N \). For example, with differentiated products, a new good may “expand the size of the market” and this may offset the effects of competition on variable profits. With heterogeneous marginal costs, the effects of competition on \( V \) are also more difficult to describe.

The fact that even in the symmetric case there are a multitude of factors that affect \( N \) is useful because it suggests that in practice information on \( N \) alone will be insufficient to identify separately all behavioral, demand and cost conditions that affect \( N \). In general, having more information, such as information on individual firm prices and quantities, will substantially improve what one can learn about market conditions.

2.2.1 Application: Entry in Small Markets

In a series of papers, Bresnahan and Reiss model the entry of retail and service businesses into small isolated markets in the United States.\footnote{See Bresnahan and Reiss (1988), Bresnahan and Reiss (1990) and Bresnahan and Reiss (1991b).} The goal of this work is to estimate how quickly entry appears to lower firms variable profits. They also seek to gauge the how large the fixed costs of setting up a business are relative to variable profits. To do this Bresnahan and Reiss, estimate a variety of models, including homogeneous and heterogeneous firm models. In their homogeneous firm models, the
number of firms flexibly enters variable profits. Specifically, because their “small” markets have at most a few firms, they allow \( V(\cdot) \) to fall by (arbitrary) amounts as new firms enter. While there are a variety of ways of doing this, Bresnahan and Reiss (1991b) assume variable profits have the form

\[
V(N^*_i, S_i, \theta) = S_i (\theta_1 + \sum_{k=2}^{M} \theta_k D_k + x_i \theta_{M+1})
\]  

where the \( D_k \) are zero-one variables equal to 1 if at least \( k \) firms have entered and \( \theta_{M+1} \) is a vector of parameters multiplying a vector of exogenous variables \( x \).

The multiplicative structure of \( V(N^*_i, S_i, \theta) \) in \( S_i \) can easily be rationalized following our previous examples (and assuming constant marginal costs). What is less obvious is the economic interpretation of the \( \theta \) parameters. The \( (\theta_2, ..., \theta_M) \) parameters describe how variable profits change as the number of entrants increases from 2 to \( M \). For example, \( \theta_2 \) is the change in a monopolist’s variable profits from having another firm enter. For the variable profit function to make economic sense, \( \theta_2, ..., \theta_M \) must all be less than or equal to zero, so that variable profits do not increase with entry. Under a variety of demand, cost and oligopoly conduct assumptions, one might also expect the absolute values of \( \theta_2, ..., \theta_M \) to decline with more entry. Bresnahan and Reiss say less about what the parameters in the vector \( \theta_{M+1} \) represent. The presumption is that represent the combined effects of demand and cost variables on (per capita) variable profits.

The size of the market, \( S \), is a critical variable in Bresnahan and Reiss’ studies. Without it, they could not hope to separate out variable profits from fixed costs. This is most easily seen by noting that without \( S \) in \( V, \theta_1 \) cannot be separately identified from a constant term in \( F \). In their empirical work, Bresnahan and Reiss assume \( S \) is itself an estimable linear function of market population, population in nearby areas and population growth.

Besides being interested in how \( \theta_2, ..., \theta_M \) declines with \( N \), Bresnahan and Reiss also are interested in estimating what they call “entry thresholds”: \( S^*_N \). The entry threshold \( S^*_N \) is the smallest overall market size \( S \) that would accommodate \( N \) potential entrants (assuming all other variables are held at average values). That is, for given \( N \) and fixed costs \( \bar{F} \), \( S^*_N = \bar{F}/V(N) \). Since \( S \) is overall market size, and larger markets are obviously needed to support more firms, it is useful to standardize \( S \) in

---

5Bresnahan and Reiss have extended their models to allowing \( F \) to vary with the number of entrants. They also explore whether profits are linear in \( S \).
order to gauge how much additional population (or whatever the units of $S$) is needed to support a next entrant. One such measure is the fraction of the overall market $S$ that a firm requires to (just) stay in the market. In the homogenous firm case this is captured by the “per-firm” threshold is $s_N = \frac{S_N}{N}$ [6]. These population thresholds can then be compared to see whether firms require increasing or decreasing numbers of customers to remain in a market as $N$ increases. Alternatively, since the units of $s_N$ may be hard to interpret, Bresnahan and Reiss recommend constructing entry threshold ratios such as $S_{N+1}/S_N$.

To appreciate what per-firm entry thresholds or entry threshold ratios reveal about demand, costs and competition, it is useful to consider the relationship between the monopoly entry threshold and per-firm thresholds for two or more firms. Casual intuition suggests that if it takes a market with 1,000 customers to support a single firm that it should take around 2,000 customers to support two firms. In other words that the per-firm entry thresholds are around 1,000 and the entry-threshold ratios are close to one. Indeed, in the homogeneous good and potential entrant case, it is not difficult to show that the entry threshold ratios will be one in competitive and collusive markets. Suppose, however, that we found that it took 10,000 customers to support a second firm (or that the entry threshold ratio was 10). What would we conclude? If we were sure that that firms’ products and technologies were roughly the same, we might suspect that the first firm was able to forestall the entry of the second competitor. But just how large is an entry threshold ratio of 10? The answer is we do not know unless we make further assumptions. Bresnahan and Reiss (1991b) provide some benchmark calculations to illustrate potential ranges for the entry threshold ratios. Returning to the Cournot example profit function (15) with $d = 0$, we would find $S_{N+1}/S_N = \frac{(N+2)^2}{(N+1)^2}$. Thus, the entry threshold ratio under these assumptions is a convex function of $N$, declining from 2.25 (duopoly/monopoly) to 1.

As we have emphasized previously, to the extent that additional data, such as prices or quantities, are available it may be possible to supplement the information that entry thresholds provide. Additionally, such information can help evaluate the validity of any maintained assumptions. For example, it may not be reasonable to assume potential entrants and their products are the same or that all entrants have the same fixed costs. Similarly, some firms might be able to price discriminate better or intertemporal considerations may play a role.

---

[6] Another way of understanding this standardization is to observe that the $N$th firm just breaks even when $V(N) S = F$. Thus, $s_n = F/(N V(N))$.  

14
Bresnahan and Reiss (1991b) argue on a priori grounds that firms fixed costs are likely to be nearly the same and that their entry threshold rations thus reveal something about competition and fixed costs. Table 1 revisits their estimates of these ratios for various retail categories. Recalling the contrast between the Cournot and perfectly collusive and competitive examples above, here we see that the ratios fall toward one as the number of entrants increases. The ratios are generally small and they decline dramatically when moving from one to two doctors, tire dealers or dentists. Plumbers are the closest industry to the extremes of perfect competition or coordination. Absent more information, Bresnahan and Reiss cannot distinguish between these two dramatically different possibilities.

| Per Firm Entry Thresholds from Bresnahan and Reiss, 1991 Table 5 |
|-----------------------|---------|---------|---------|---------|
| Profession           | $S_2/S_1$ | $S_3/S_2$ | $S_4/S_3$ | $S_5/S_4$ |
| Doctors              | 1.98     | 1.10     | 1.00     | 0.95     |
| Dentists             | 1.78     | 0.79     | 0.97     | 0.94     |
| Plumbers             | 1.06     | 1.00     | 1.02     | 0.96     |
| Tire Dealers         | 1.81     | 1.28     | 1.04     | 1.03     |

In an effort to understand the information in entry thresholds, Bresnahan and Reiss (1991b) also collected information on the prices of standard tires from tire dealers in both small and large markets. They then compared these prices to their entry threshold estimates. Consistent with Table 1, tire dealers’ prices did seem to fall with the first few entrants; they then leveled off after five entrants. Curiously, however, when they compared these prices to those in urban areas they found that prices had in some cases leveled off substantially above those in urban areas where there are presumably a very large number of competitors.

2.3 Observables and Unobservables

So far we have focused on deriving how observables such as $x$, $S$, $N$ and $P$ affect entry decisions, and said little about how assumptions about unobservables affect estimation. Already we have seen that empirical models of market structure are likely to rest heavily on distributional assumptions. This subsection considers what types of economic assumptions might support these assumptions.
To derive the stochastic distribution of unobserved profits, we can proceed in one of two ways. One is to make assumptions about the distribution of underlying demand and costs. From these distributions and a model of firm behavior, we can derive the distribution of firms’ unobserved profits. The second way is to simply assume distributions for variable profits and fixed costs that appear economically plausible and yet are computationally tractable. The strength of the first of these approaches is that it makes clear how unobserved demand and cost conditions affect firm profitability and entry; a disadvantage of this approach, which is anticipated in the second approach, is that it can lead to empirically intractable models.

To implement the first approach, we must impose specific functional forms for demand and cost. Suppose, for example, the researcher observes inverse market demand up to an additive error and unknown coefficients $\theta^d$

$$P = D(x, Q, \theta^d) + \epsilon^d$$

and total costs are linear in output

$$TC(q) = F(w) + \epsilon^F + (c(w, \theta^c) + \epsilon^c)q.$$  

In these equations, the firm observes the demand and cost unobservables $\epsilon^d$ and $\epsilon^c$, the $w$ are $x$ are exogenous variables and $q$ is firm output. Suppose in addition firms are symmetric and each firm equates its marginal revenue to marginal cost. The researcher then can calculate the “mark-up” equation:

$$P = b(x, q, Q, \theta^d) + c(w, \theta^c) + \epsilon^m.$$  

Here, $b$ is the derivative of $D$ with respect to firm output multiplied by $q$ and $\epsilon^m = \epsilon^d - \epsilon^c$. Notice that because we assumed the demand and cost errors are additive, they do not enter the $b(\cdot)$ and $c(\cdot)$ directly. (The errors may enter $b(\cdot)$ indirectly if the firms’ output decisions depend on the demand and cost errors.)

This additive error structure has proven particularly convenient in the market power literature (see Bresnahan (1989)), since it permits the researcher to employ standard instrumental variable or generalized method of moment techniques to estimate demand and cost parameters from observations on price and quantity. This error structure, however, complicates estimation methods based on the number of firms. To see this, return to the profit function expression (15) in the linear demand example. If we add $\epsilon_d$ to the demand intercept and $\epsilon_c$ to marginal cost we obtain

$$\pi_i(N_i^*, S_i) = V(N_i^*, S_i, \theta) - F_i = (\theta_1 + \epsilon_m)S_i \frac{(1 + \theta_2 S_i^2)}{(N_i^* + 1 + 2\theta_2 S_i)^2} - F_i - \epsilon^F,$$
That is, profits are linear in the fixed cost error but quadratic in the demand and marginal cost errors. Consequently, if we assumed the demand and cost errors were i.i.d., profits would be independent but not identically distributed across markets that varied in size \((S)\).

It should perhaps not be too surprising that the reduced form distribution of firms’ unobserved profits can be a nonlinear function of unobserved demand and cost variables. While these nonlinearities complicate estimation, they do not necessarily preclude it. For example, to take the above profit specification to data, one might assume the fixed costs have an additive normal or logit error, or a multiplicative log-normal error. These assumptions lead to tractable expressions for the likelihood function expressions (such as (6)) conditional on values of \(\epsilon^d\) and \(\epsilon^c\). The researcher could then in principle attempt estimation by integrating out the demand and cost errors using either numerical methods or simulation techniques.

2.4 Demand, Supply and Endogenous \(N\)

So far we have only considered models based on the number of firms in a market, and not price or quantity. In some applications, researchers are fortunate enough to have price and quantity information in addition to information on market structure. This subsection asks what the researcher gains by modeling market structure in addition to price and quantity.

It might seem at first that there is little additional value to modeling market structure. Following the literature on estimating market power in homogeneous product markets, one could use price and quantity information alone to estimate industry demand and supply (or markup) equations such as:

\[
Q = Q(P, X, \theta^d, \epsilon^d)
\]

and

\[
P = P(q, N, W, \theta^c, \epsilon^c).
\]

In these equations, \(Q\) denotes industry quantity, \(q\) denotes per firm (symmetric) quantity, \(P\) is price, \(X\) and \(W\) are exogenous demand and cost variables, and \(\epsilon^d\) and \(\epsilon^c\) are demand and supply unobservables. The parameter vectors \(\theta^d\) and \(\theta^c\) represent industry demand and cost parameters, such as those found in the previous subsection.

\[\text{To our knowledge this approach has not been attempted. This is perhaps because a proof that such an approach would work and its econometric properties remain to be explored.}\]
Provided $X$ and $W$ contain valid instruments for price and quantity, it would appear an easy matter to use instrumental variables to estimate $\theta^d$ and $\theta^c$. Thus, the only benefit to modeling model market structure would seem to be that it allows the researcher to estimate fixed costs (which do not enter the demand and supply equations above). This impression overlooks the fact that $N$ (or some other measure of industry concentration) may appear separately in the supply equation and thus require instruments.

The examples in previous subsections illustrate why the endogeneity of $N$ can introduce complications for estimating $\theta^d$ and $\theta^c$. They also suggest potential solutions and instruments. In previous examples, the number of firms $N$ was determined by a threshold condition on firm profits. This threshold condition depends (nonlinearly) on the exogenous demand ($x$) and variable cost ($w$) variables and the demand and total cost unobservables that make up the demand and supply errors. Thus, to estimate the parameters of demand and supply equations consistently, we have to worry about finding valid instruments for the number of firms. The most compelling instruments for the number of firms (or other market concentration measures) would be exogenous variables that affect the number of firms but not demand or supply. One such source are observables that only enter fixed costs. Examples might include the prices of fixed factors of production or measures of opportunity costs.

In some applications, it may be hard to come by exogenous variables that affect fixed costs, but not demand and variable costs. In such cases, functional form or error term restrictions might justify a particular choice of instrument or estimation method. For instance, in the linear demand and marginal cost example of subsection 2.4 if we assume $d = 0$ (constant marginal costs), then we can use market size $S$ as an instrument for the number of firms. This is essentially the logic of Bresnahan and Reiss, who note for the markets they study that market size is highly correlated with the number of firms in a market. Market size also is used explicitly as an instrument in Berry and Waldfogel (1999).

2.4.1 Application: Market Structure and Competition in Radio

Berry and Waldfogel (1999) examine the theoretical hypothesis that entry can be socially inefficient. They do this by comparing advertising prices, listening shares and numbers of stations broadcasting in different radio markets. Specifically, they

---

\[ ^8 \text{Notice that per capita total quantity } Q \text{ and per capita firm quantity } q \text{ are independent of } S. \text{ Thus, } S \text{ does not enter the per capita demand function or the firm's supply equation.} \]
ask whether the fixed costs of entry exceed the social benefits of new programming (greater listening) and more competitive advertising prices.

To compute private and social returns to entry, Berry and Waldfogel must estimate (i) the fixed costs of entrants; (2) by how much new stations expand listening; and (3) by how much entry changes advertising prices. They do all this by developing an empirical model in which homogeneous stations “produce” listeners and then “sell” them to advertisers. The economic primitives of the model include: a listener choice function; an advertiser demand function; and a specification for station fixed costs.

Berry and Waldfogel model radio listeners within a market as having correlated extreme value preferences for homogeneous stations; an outside good (not listening) also is included. Under their station homogeneity and stochastic assumptions, what varies across markets is the fraction of listeners and non-listeners, which are in turn affected by the number of entrants. Specifically, under their assumptions, listener \( L_i \) relative to non-listener \((1 - L_i)\) shares in market \(i\) are related by

\[
\ln\left(\frac{L_i}{1 - L_i}\right) = x_i\beta + (1 - \sigma) \ln(N_i) + \xi_i. 
\]

That is, the odds for listening depend on a set of market demographics, \(x_i\), the number of (homogeneous) stations in the market, \(N_i\), and a market-specific unobservable, \(\xi_i\). The parameter \(\sigma\) controls the correlation of consumers’ idiosyncratic preferences for stations. When \(\sigma = 1\), consumers’ unobserved preferences for the homogeneous stations are perfectly correlated and thus the entry of new stations does not expand the number of listeners. When \(\sigma = 0\), as in the case in a conventional logit model, the entry of an otherwise identical station expands the size of the market because some consumers will have an idiosyncratic preference for it (relative to other stations and not listening).

As a demand equation, (22) is linear in its parameters and thus easily estimated by linear estimation techniques. As we pointed out in the beginning of this subsection, having \(N\) on the right hand side poses a problem – the number of radio stations in a market, \(N_i\), will be correlated with the market demand unobservable \(\xi\). Thus, Berry and Waldfogel must find an instrument for \(N_i\). For the reasons outlined earlier, the size of the radio market provides a good instrument for the number of stations. It does not enter consumer preferences directly and yet is something that affects the overall size of the market.

Rysman (2004) studies the welfare effects of entry in the market for telephone Yellow Pages and also considers possible network effects.
Next, Berry and Waldfogel introduce advertisers’ demand for station listeners. Specifically, they assume that demand has the constant elasticity form:

$$\ln(p_i) = x_i \gamma - \eta \ln(L_i) + \omega_i,$$

(23)

where $p_i$ is the price of advertising, and $\omega$ is the demand error. Once again, market size is a good instrument for listening demand, which may be endogenous. Together, the listening share and pricing equations give the revenue function of the firm.

Since the marginal cost of an additional listener is literally zero, Berry and Waldfogel model all costs as fixed costs. The fixed costs must be estimated from the entry equation. As in our earlier discussions, with a homogeneous product and identical firms, $N_i$ firms will enter if

$$R(N_i + 1, x_i, \theta) < F_i < R(N_i, x_i, \theta),$$

(24)

where $R(\cdot)$ is a revenue function equal to $p_i(N_i, \gamma, \eta) M_i L_i(N_i, \beta, \sigma)$ (where $M_i$ is the population of market $i$). Thus, with appropriate assumptions about the distribution of unobserved costs and revenues across markets, Berry and Waldfogel can use ordered dependent variables techniques to learn the distribution of $F$ across markets. Knowing this distribution, Berry and Waldfogel compare the welfare effects of different entry regimes: free entry, socially optimal entry and monopolized radio spectrum. An important problem they face is that the market participants are the radio stations and the advertisers; the benefit to listeners is an unpriced external benefit. Looking only at market participants, there appears to be too much entry relative to the social optimum. This is because in many markets, the incremental station generates a small number of the additional listeners that advertisers value. However, the external benefit to listeners could easily offset this finding.

3 Firm Heterogeneity

We have so far explored models with identical firms. In reality, firms have different costs, sell different products, and occupy different locations. It is therefore important to explore how entrant heterogeneities might affect estimation. Firm heterogeneities can be introduced in a variety of ways, including observed and unobserved differences.

More realistically, this fixed cost would be endogenously chosen and would affect the quality of the station, which Berry and Waldfogel do not model.
in: firms’ fixed and variable costs, product attributes, and product distribution. A
first important issue to consider is how these differences arose. In some cases, the
differences might reasonably be taken as given and outside the firms’ control. In other
cases, such as product quality, the differences are under a firm’s control and thus will
have have to be modeled along with market structure. Almost all empirical models
to date have adopted the approach that differences among firms are given. Although
we too adopt this approach in much of what follows, the modeling of heterogeneities
is ripe for exploration.

As we shall see shortly, empirical models with heterogeneous potential entrants
pose thorny conceptual and practical problems for empirical researchers. Chief among
them are the possibility that entry models can have multiple equilibria, or worse, no
pure-strategy equilibria. In such cases, standard approaches to estimating parameters
may break down, and indeed key parameters may not longer be identified.

Although we did not note these problems in our discussion of homogeneous firm
models, they can occur there as well. We discuss them here because they pose easier
to grasp problems for researchers trying to match firm identities or characteristics to
a model’s predictions about who will enter. As noted by Sutton (2006) and others,
the problems of nonexistence and nonuniqueness in theory and empirical models have
traditionally been treated as nuisances – something to be eliminated by assumption
if at all possible. We will provide several different examples of this approach in this
section. We should remark, however, that multiplicity or nonexistence issues may
be a fact of markets. For this reason we will consider alternative solutions in the
following section.

In this section we emphasize that heterogeneous firm entry models differ along
two important dimensions: (1) the extent to which heterogeneities are observable or
unobservable to the econometrician; and (2) the extent to which firms are assumed
to be uncertain about the actions or payoffs of other firms. Both of these dimensions
critically affect the identification and estimation of entry models. For example, McK-
elvey and Palfrey (1995) and Seim (2004) have shown how introducing asymmetric
information about payoffs can mitigate multiple equilibrium problems. Others (e.g.,
Bresnahan and Reiss (1990), Berry (1992) and Mazzeo (2002)) have explored how ob-
serving the timing of firms’ decisions can eliminate nonexistence and nonuniqueness
problems.

To illustrate these economic and econometric issues and solutions, we begin with
the simplest kind of heterogeneity – heterogeneity in unobserved fixed costs. We first
discuss problems that can arise in models in which this heterogeneity is known to the
firms but not the researcher. We also discuss possible solutions. We then discuss how entry models change when firms have imperfect information about their differences.

3.1 Complications in Models with Unobserved Heterogeneity

To start, consider a one-play, two-by-two entry game in which there are two potential entrants, each with two potential strategies. Suppose firms 1 and 2 have perfect information about each other and earn $\pi_1(D_1, D_2)$ and $\pi_2(D_1, D_2)$ respectively from taking actions $(D_1, D_2)$, where an action is either 0 (“Do Not Enter”) or 1 (“Enter”).

Following our earlier derivations, we would like to derive equilibrium conditions linking the firms’ observed actions to inequalities on their profits. A natural starting point is to examine what happens when the firms are simultaneous Nash competitors – that is, they make their entry decisions simultaneously and independently. Additionally, we assume that entry by a competitor reduces profits and that a firm earns zero profits if it does not enter. Under these conditions, the threshold conditions supporting the possible entry outcomes are:

<table>
<thead>
<tr>
<th>Market Outcome</th>
<th>N</th>
<th>Conditions on Profits</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Firms</td>
<td>0</td>
<td>$\pi_1^M &lt; 0$</td>
</tr>
<tr>
<td>Firm 1 Monopoly</td>
<td>1</td>
<td>$\pi_1^M &gt; 0$</td>
</tr>
<tr>
<td>Firm 2 Monopoly</td>
<td>1</td>
<td>$\pi_1^D &lt; 0$</td>
</tr>
<tr>
<td>Duopoly</td>
<td>2</td>
<td>$\pi_2^D &gt; 0$</td>
</tr>
</tbody>
</table>

Here, the notation $\pi_j^M$ and $\pi_j^D$ denote the profits firm $i$ earns as a monopolist and duopolist respectively.

Following our earlier discussions, the researcher would like to use observations on $(D_1, D_2)$ to recover information about the $\pi_j^M$ and $\pi_j^D$. To do this, we have to specify how firms’ profits differ in observable and unobservable ways. In what follows we decompose firms’ profits into an observable (or estimable) component and an

---

11Bresnahan and Reiss (1991a) discuss the significance of these assumptions.
additively separable unobserved component. Specifically, we assume

\[ \pi_j = \begin{cases} 
0 & \text{if } D_j = 0 \\
\pi_j^M(x, z_j) + \epsilon_j & \text{if } D_j = 1 \text{ and } D_k = 0 \\
\pi_j^D(x, z_j) + \epsilon_j & \text{if } D_j = 1 \text{ and } D_k = 1 
\end{cases} \]

In these equations, the \( \bar{\pi} \) terms represent observable profits. These profits are functions of observable market \( x \) and firm-specific \( z_j \) variables. The \( \epsilon_j \) represent profits that are known to the firms but not to the researcher. Notice that this additive specification presumes that competitor \( k' \)'s action only affects competitor \( j \)'s profits through observed profits; \( k' \)'s action does not affect that part of profits the researcher cannot observe. This special assumption simplifies the analysis. One rationale for it is that the error \( \epsilon_j \) represents firm \( j \)'s unobservable fixed costs, and competitor \( k \) is unable to raise or lower their rival’s fixed costs by being in or out of the market.

The restrictions in Table 2, along with assumptions about the distribution of the \( \epsilon_j \), link the observed market structures to information about firms’ demands and costs. Figure 1 displays the values of firms’ monopoly profits that lead to the four distinct entry outcomes. Following the rows of Table 2, the white area to the southwest represents the region where both firms’ monopoly profits are less than zero and neither firm enters. Firm 1 has a monopoly in the area to the southeast with horizontal gray stripes. There, firm 1’s monopoly profits are positive and yet firm 2’s duopoly profits (by assumption less than monopoly profits) are still negative. Similarly, firm 2 has a monopoly in the northeast area with vertical gray stripes. There, firm 2’s monopoly profits are positive and firm 1’s duopoly profits negative. Finally, the solid gray region to the northeast corresponds to the last row of Table 2 in which both firms enter.

The shading of the figure shows that given our assumptions there always is at least one pure-strategy Nash equilibrium. The center cross-hatched, however, supports two pure-strategy equilibria: one in which firm 1 is a monopolist and one where firm \( w \) is a monopolist. Absent more information, the conditions in Table 2 do not provide an unambiguous mapping from equilibria to inequalities on profits. This causes problems for constructing likelihood functions (see Bresnahan and Reiss (1991)). In a moment, we will discuss potential fixes for this problem.

Besides illustrating what problems can arise when relating discrete entry outcomes
to equilibrium conditions on profits, Figure 1 also shows why simple probit or logit models are inadequate for modeling heterogeneous firms’ entry decisions. A simple probit would presume \( j \) would enter whenever \( \bar{\pi}_j > -\epsilon_j \). Notice, however, that in the regions to the north and to the east of the center cross-hatched rectangle that although the monopoly profits of both firms are positive, the firm with positive duopoly profits can preclude the other from being a monopolist. That is, what the standard probit model does not recognize is the possibility of pre-emption – that the entry of one firm could cause the other firm to exit the market.\(^{12}\) A properly specified model of simultaneous entry decisions needs to recognize this interdependence of profits.

Finally, we have thus far focused on what happens when there are two potential entrants. The points made in this duopoly example continue to hold as one moves to larger concentrated oligopolies. In general, the researcher will have conditions that relate firms’ latent profits to their discrete decisions. To illustrate, a Nash equilibrium \( D^*\{D^*_1, ..., D^*_N\} \) requires

\[
D^* \cdot \pi(D^*) > 0 \quad (1 - D^*) \cdot \pi(D^* + S_j \cdot (1 - D^*)) \leq 0
\]

for all \( S_j \), where \( \pi \) is an \( N \)-vector of firm profit functions, \( \cdot \) is element-by-element multiplication, and \( S_j \) is a unit vector with a one in the \( j \)th position. The first condition requires that all entrants found it profitable to enter. The second condition requires that no potential finds it profitable to enter.\(^{13}\) This extended definition again does not rule out multiple equilibria. In general, if several firms have similar \( \epsilon \)’s then there may be values of firm profits (such as in the center of Figure 1) which would simultaneously support different subsets of firms entering.

### 3.2 Potential Solutions to Multiplicity

A variety of authors have proposed solutions to the multiplicity problem, beginning with Bjorn and Vuong (1984), Bresnahan and Reiss (1991a), Bresnahan and Reiss (1991b) and Berry (1992). These solutions include (individually or in combination) changing what is analyzed, changing the economic structure of the underlying game, and changing assumptions about firm heterogeneities.

\(^{12}\)In Figure 1 this occurs to the north and east of the cross-hatched box. There, firm \( j \)’s monopoly profits are positive but firm \( j \)’s duopoly profits are negative and firm \( k \)’s duopoly profits are positive. Firm \( k \) can therefore preempt firm \( j \).

\(^{13}\)Because entry reduces competitor profits, if the second condition holds for all potential entrants individually, it holds for all combinations of potential entrants.
One strategy is to consider only the probabilities of aggregated outcomes that are robust to the multiplicity of equilibria. Another common strategy is to simply place additional conditions on the model that guarantee a unique equilibrium. One econometric strategy is to model the probability of each outcome and then include in the estimation procedure additional parameters (or additional unknown non-parametric functions). Finally, a recently-proposed alternative is to accept that some models with multiple equilibria will not make precise predictions about the probability of various events, but to note that such models can still place useful restrictions on model parameters. We consider examples of each of these approaches in the following subsections.\footnote{Sweeting (2005) notes that there may be cases where the existence of multiple equilibrium actually helps in estimation. The reason is that multiple equilibrium can create variance in data that otherwise would not be present and this variance can potentially help to estimate a model.}

### 3.2.1 Aggregating Outcomes

Bresnahan and Reiss (1988) and Bresnahan and Reiss (1991a) observe that although the threshold inequalities in Table 2 describing firms’ decisions are not mutually exclusive, the inequalities describing the number of firms are mutually exclusive. In other words, the model uniquely predicts the number of firms that will enter, but not their identities. To see this, return to Figure 1. There, the number of firms is described by the following mutually exclusive and exhaustive regions: the white region (no firms), the solid gray region (duopoly) and the region with gray lines (monopoly). Given assumptions about the distribution of firms’ profits, it would now be possible to write down a likelihood function for the profit parameters.

While changing the focus from analyzing individual firm decisions to a single market outcome ($N$) can solve the multiplicity problem, it is not without its costs. One potential cost is the loss of information about firm heterogeneities. In particular, it may no longer be possible to identify all the parameters of individual firms’ observed and unobserved profits from observations on the total number of firms than entered.\footnote{See, for example, Bresnahan and Reiss (1991a) and Andrews, Berry, and Jia (2004).}

### 3.2.2 Timing: Sequential Entry with Predetermined Orders

An alternative response to multiple equilibria in perfect-information, simultaneous-move entry games is to assume that firms instead make decisions sequentially. While
this change is conceptually appealing because it guarantees a unique equilibrium, it may not be practically appealing because it requires additional information or assumptions. For instance, the researcher either must: know the order in which firms move; make assumptions that permit the order in which firms move to be recovered from the estimation; or otherwise place restrictions on firms profit functions or the markets firms can enter. We discuss each of these possibilities in turn.

When firms make their entry decisions in a predetermined order, it is well known that early movers can preempt subsequent potential entrants (e.g., Bresnahan and Reiss (1990) and Bresnahan and Reiss (1991a)). To see this, recall the structure of the equilibrium payoff regions of Figure 1. There, the payoffs in the center rectangle would support either firm as a monopolist. Now suppose that we knew or were willing to assume that firm 1 (exogenously) moved first. Under this ordering, the equilibrium threshold conditions are:

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Market Outcome & N & Conditions on Profits \\
\hline
No Firms & 0 & $\pi_1^M < 0$ & $\pi_2^M < 0$ \\
Firm 1 Monopoly & 1 & $\pi_1^M > 0$ & $\pi_2^D < 0$ \\
Firm 2 Monopoly & 1 & $\pi_2^M > 0$ & $\pi_1^D < 0$ \\
& & $\pi_2^D > 0$ & $\pi_1^D > 0$ \\
Duopoly & 2 & $\pi_2^D > 0$ & $\pi_1^D > 0$ \\
\hline
\end{tabular}
\caption{Two-Firm Market Structure Outcomes for an Single-Stage Sequential-Move Game}
\end{table}

The sole difference between Table 3 and Table 2 is that the region where firm 2 can be a monopolist shrinks. Specifically, by moving first, firm 1 can preempt the entry of firm 2 in the center cross-hatched area in Figure 1.

This change eliminates the multiplicity problem and leads to a coherent econometric model. If, for example, the researcher assumes the joint distribution of unobserved profits is $\phi(\cdot, x, z, \theta)$, then the researcher can calculate the probability of observing any equilibrium $D^*$ as

$$Pr(D^*) = \int_{A(D^*, x, z, \theta)} \phi(\epsilon, x, z, \theta) \, d\epsilon.$$  \hfill (25)

In this expression, $A(D^*, x, z, \theta)$ is the region of $\epsilon$’s that leads to the outcome $D^*$. For example, in Figure 1 $A(0, 1)$ would correspond to the northwest region of firm 2.
monopolies. There are two main problems researchers face in calculating the probabilities when there are more than a few firms: 1) how to find the region $A$; and 2) how to calculate the integral over that region.

The problem of finding and evaluating the region $A$ can become complicated when there are more than a few firms. Berry (1992) solves this problem via the method of simulated moments, taking random draws on the profit shocks and then solving for the unique number and identity of firms. With a sufficient number of draws (or more complicated techniques of Monte Carlo integration) it is also possible to construct a simulated maximum likelihood estimator.

3.2.3 Efficient (Profitable) Entry Models

One criticism that can be leveled against models that always assume that firms always move in a particular order is that they can imply that a very inefficient first-mover will preempt a much more efficient second-mover. While the preemption of more efficient rivals may be realistic outcome in some markets, it may not fit the realities of other markets.

An alternative modeling strategy would be assume that inefficient entry never occurs. That is, that the most profitable entrant always is able to move first. In our two-firm model, for example, we might think of there being two entrepreneurs that face the profit possibilities displayed in Figure 1. The entrepreneur who moves first is able to decide whether they will be firm 1 or firm 2 and then whether they will enter. In this case, the first entrepreneur will always elect to be the entrant with the greatest profits. This means that the center region of multiple monopoly outcomes in Figure 1 will now be divided as in Figure 2. In Figure 2, the dark, upward-sloping $45^\circ$ line now divides the monopoly outcomes. The area above the diagonal and to the northwest represents outcomes where firm 2 is more profitable than firm 1. In this region, the initial entrepreneur (the first-mover) chooses to be firm 2. Below the diagonal, the opposite occurs – the initial entrepreneur (the first-mover) chooses to be firm 1. The thresholds that would be used in estimation are thus

Figure 2 About Here

Table 5
Two-Firm Market Structure Outcomes for an
3.2.4 Estimating the Probabilities of Different Equilibria.

Another possible approach to multiplicity is to assume that the players move sequentially, but to treat the order as unknown. Although the sequential assumption guarantees that each entry game did have a unique outcome, it does not guarantee that the researcher has a complete model of this outcome. For instance, absent further assumptions, even though entry is sequential the researcher’s econometric model will again mirror the multiplicity of the simultaneous-move model.

One response to this problem is to add a mechanism to the entry model that dictates the order in which the players move. Indeed, the assumption that the order is the same (or known) across all markets is a trivial example of a mechanism. One alternative mechanism to assume that the order is randomly determined. Such a

This two-stage model of sequential has several advantages and disadvantages when compared to previous models. On the positive side it resolves the multiplicity problem and the need to observe which firm moved first. For example, if we observe a firm 2 monopoly we know that the first-mover chose to be firm 2 because this was the more profitable of the two monopolies. Yet another potential advantage of this model is that in cases where the researcher observes which duopolist entered first, the researcher has additional information that potentially may result in more precise parameter estimates.

With these advantages come disadvantages however. Chief among them is a computational disadvantage. This disadvantage can be seen in the non-rectangular shape of the monopoly outcome regions. These non-rectangular shapes can considerably complicate estimation – particularly if the firms’ unobserved profits are assumed to be correlated. Again, Berry (1992) solves this problem by simulating the market outcome, which is this case is trivial to calculate even when the regions of integration are difficult.
possibility was explored by Bjorn and Vuong (1984). One version of their approach would for example assign a probability \( \lambda \) to one of the two monopolies occurring. The researcher would then attempt to estimate this probability along with the other parameters.

Tamer (2003) extends this approach to let the probability of each equilibria depend on the exogenous \( x \)’s observed in the data. In the two-firm case, he introduces an unknown function \( H(x) \), which is the probability that firm 1 enters when the draws place us in the region of multiple equilibria. (More generally, one could let that probability depend on the unobservable as well, giving a new function \( H(x, \epsilon) \).) Tamer estimates \( H \) as an unknown non-parametric function.

Tamer (2003) notes that, under suitable assumptions, the heterogeneous firm entry model is identified simply from information on the (uniquely determined) number of entering firms. However, adding an explicit probability for each equilibria can increase the efficiency of the MLE estimates.

There are several potential shortcomings of this approach. At a practical level, depending upon how this probability is specified and introduced into the estimation, the researcher may or may not be able to estimate these probabilities consistently. Additionally, once there are more than two potential entrants, the number of payoff regions that can support multiple outcomes can proliferate quickly, necessitating the introduction of many additional probability parameters to account for the frequency with which each outcome occurs. On a more general level, there is the conceptual modeling question of how to interpret these probabilities. Is the researcher really willing to entertain that there is a random selection process?

3.2.5 A Bounds Approach.

In a variety of contexts, Manski suggests the possibility of a “bounds” approach to estimation in cases where a model is somewhat incomplete and therefore does not make clear predictions about outcomes in the data.\(^{16}\) The case of multiple equilibria seems to fit this case nicely.

The basic idea is that while the model may not make exact predictions about outcomes, it does still restrict the range of possible outcomes. In some highly parameterized cases, these restrictions may still serve to point-identify the model. In many cases, however, the qualitative restrictions implied by the model with multiple

\(^{16}\)See, for example, Manski (1995) and Manski and Tamer (2002).
equilibria may identify, for any particular data-generating process, a non-trivial set of parameters rather than a single point. One of Manski’s contributions is to point out that “set identification” is often useful in either [i] testing particular hypotheses or else [ii] illustrating the range of outcomes (of, say, a proposed policy) that are consistent with the model and the data.

In the context of oligopoly entry models, Ciliberto and Tamer (2003) and An Explanation of Anomalous Behavior in Binary-Choice Games: Entry, Voting, Public Goods, and the Volunteers’ Dilemma Andrews, Berry, and Jia (2004) consider the problem of placing bounds on the parameters of entry models. We can illustrate the idea in a simple context.

Suppose that the profit for firm $j$ of entering into market $m$ is

$$\bar{\pi}(y_{-j,m}, x_{jm}, \theta) + \epsilon_{jm},$$

(26)

$y_{-j,m}$ is a vector of dummy variables indicating whether the firm’s rivals have entered, $x_{jm}$ is a vector of profit shifters, $\theta$ is a vector of parameters to be estimated and $\epsilon_{jm}$ is an unobserved profit shifter. Here, vector $y$ describes the equilibrium market structure, with elements of $y$ equal to 1 for those firms that enter and equal to 0 otherwise.

As before, if the firm enters, the best reply condition must be satisfied:

$$\bar{\pi}(y_{-j,m}, x_{jm}, \theta) + \epsilon_j \geq 0,$$

(27)

and if the firm does not enter, then

$$\bar{\pi}(y_{-j,m}, x_{jm}, \theta) + \epsilon_j \leq 0.$$

(28)

In the case of multiple equilibria, these conditions are necessary but not sufficient, because the existence of multiple equilibria means that the same vectors of $(x, \epsilon)$ might also lead to another outcome, $y'$.

To outline the method, note that we can calculate the joint probability (across firms within a market) that the best reply conditions in (27) and (28) hold. This probability is not the probability of the observed entry choices, but it is an upper bound on that probability. This follows from the fact that necessary conditions are weaker than necessary-and-sufficient conditions and so the probability that the necessary conditions hold is greater than or equal to the probability of the observed entry decisions.
To be more formal, say that in market $m$ we observe the market structure outcome $y_m$ and the exogenous shifters $x_m$. For a given parameter $\theta$, we can begin by defining $\Omega(y, x, \theta)$, the set of $\epsilon$’s that jointly satisfy the necessary conditions in (27) - (28) for a pure-strategy Nash equilibrium. Define the true probability of the outcome $y$, at the true parameter $\theta_0$ as $P_0(y | x)$. By the definition of a necessary condition, then, at the true parameter vector the probability of the necessary condition must weakly exceed the probability of the equilibrium event:

$$Pr(\epsilon \in \Omega(y, x, \theta_0) \geq P_0(y | x))$$

(29)

At best, then, we know that a candidate parameter $\theta$ is consistent with the model and the data generating process only if

$$Pr(\epsilon \in \Omega(y, x, \theta) \geq P_0(y | x))$$

(30)

The “identified set” of parameters is then the set of $\theta$’s that satisfy (30). Our intuition is that a set of inequality restrictions will typically not identify a single point, but for a given parameterization it may be the case that only a single $\theta$ will satisfy (30). Or, the model may be rejected, with no $\theta$ that satisfies An Explanation of Anomalous Behavior in Binary-Choice Games: Entry, Voting, Public Goods, and the Volunteers’ Dilemma.

Even when we have only “set identification”, that set may allow us to reject interesting hypotheses, such as the hypothesis that two firms do not compete with each other. Also, the set of identified parameters will typically place bounds on the outcomes of policy experiments, such as questions about how market structure would change with the number of potential entrants.

These methods raise the interesting econometric question of how to form a good sample analog of (30) and how to prove consistency of the resulting estimator. Even more difficult is the question of how to place a confidence region on the set of parameters that satisfy the model. This is very much on-going research. One of the first papers to address the inference problem in a general way is Chernozhukov, Hong, and Tamer (2002) and there is also on-going research in Andrews, Berry, and Jia (2004) and in Ciliberto and Tamer (2003).


In the next subsection, we consider two empirical applications that rely on more traditional econometric methods.
3.3 Applications with Multiple Equilibria

3.3.1 Application: Motel Entry

Bresnahan and Reiss (1991) discuss a variety of approaches to modeling discrete games, including entry games. Several papers have adopted their structure to model firms’ choices of discrete product types and entry decisions. Many of these papers also recognize the potential for multiple equilibria, and many adopt solutions that parallel those discussed above.

Mazzeo (2002), for example, models the entry decisions and quality choices of motels that locate near highway exits. Specifically, he has data on the number of high and low quality motels at a number of geographically distinct set of highway exits. Thus, his discrete game is one where each potential entrant chooses a quality and whether to enter. It should be immediately clear that this is a setting where non-uniqueness and non-existence problems can easily arise. For example, consider the following table listing the (assumed symmetric) profit opportunities for high (H) and low (L) quality hotels as a function of the number of entrants of each type, NL and NH.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
N_L & 0 & 1 & 2 & 3 \\
\hline
0 & 0, 0 & 0, 4 & 0, 1.5 & 0, -1 \\
1 & 3, 0 & 2, 2 & -1, -1 & -2, -1 \\
2 & 1, 0 & -1, -1 & -2, -2 & -4, -3 \\
3 & -1, 0 & -2, -3 & -3, -4 & -5, -5 \\
\hline
\end{tabular}
\caption{Entrant Profits for Qualities L and H}
\end{table}

We note that the firms’ profits decline as more firms of either type enter the local market. It is easy to verify that there profit outcomes support three Nash simultaneous-move equilibria: (2,0), (1,1), and (0, 2). Each of these outcomes results in two entrants.

Mazzeo recognizes the potential for multiple outcomes such as these and explores
two responses. The first is to assume that these firms move sequentially. This assumption will guarantee a unique prediction for the game. Moreover, because here the firms are ex ante symmetric, it may also possible to estimate parameters of the profit functions even when one does not know the exact order of entry. For example, in the above example, it is clear that (1,1) is the unique sequential-move equilibrium provided we assume that the entry of a same-quality duopolist lowers profits more than a different-quality monopolist. With this assumption, it does not matter whether the first mover selected high or low quality, the second mover will always find it optimal to pick the other quality.

As we have noted previously, in some cases the fact that firms move sequentially may advantage or disadvantage certain firms. For instance, ex post a firm may wish they could change their decision. Consider, for example, the following profit tableau:

<table>
<thead>
<tr>
<th>( N_L )</th>
<th>( N_H )</th>
<th>( \pi_L(N_L, N_H) )</th>
<th>( \pi_H(N_L, N_H) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5, 4</td>
<td>3, 3</td>
<td>1, 1.5</td>
</tr>
<tr>
<td>2</td>
<td>3, 3</td>
<td>2, 2</td>
<td>-1, 1</td>
</tr>
<tr>
<td>3</td>
<td>-1, 0</td>
<td>-2, -1</td>
<td>-3, -4</td>
</tr>
</tbody>
</table>

Here there are two simultaneous-move Nash equilibria: (2, 2) and (1, 3). The unique sequential-move equilibrium is (1, 3), as a third high-quality firm would always choose to enter if two low-quality firms were in the market.

Perhaps because a sequential-move equilibrium can result in inefficient outcomes, Mazzeo also considers two-stage equilibria where firms first make their entry decisions simultaneously and then make their quality decisions simultaneously. The equilibrium of this two-stage game in Table 6a is (2, 2), as now four firms initially commit to enter and then they split themselves equally among the qualities. This staging of the game here in effect selects the efficient entry outcome.

Because in general this second equilibrium concept need not result in unique entry and quality outcomes, Mazzeo must place additional structure on firms’ observed and unobserved payoffs. In his empirical work, he assumes firm \( j \) of quality type \( k \) in
market $i$ has profits

$$\pi_{ji} = x_i \beta_k + g_k(N_{Li}, N_{Hi}, \theta_k) + \epsilon_{ki}. \quad (31)$$

where $N_L$ and $N_H$ are the numbers of high and low quality competitors. It is important to note that both the observable and unobservable portion of firm profits have no firm-level idiosyncracies. In other words, profits are the same for all firms of a given quality in a given market. This assumption appears to be required because without it the specific order in which firms moved in a given market might alter the market structure outcome. This assumption also considerably simplifies estimation.

The function $g(\cdot)$ is introduced to allow the number of competitors of either quality type to affect variable profits. Following Bresnahan and Reiss (1988), Mazzeo allows considerable flexibility in $g(\cdot)$ by using a set of dummy variables to shift $g(\cdot)$ with $N_L$ and $N_H$. A key restriction, however, is that profits must decline faster with the entry of a firm of the same quality than the entry of a firm with a different quality. While this assumption seems reasonable in his particular application, it may not in others when the effect of entry depends more on factors idiosyncratic to the entrant.\footnote{For instance,}

In his empirical analysis, Mazzeo finds that the restrictions he must place of the profit functions of high and low quality firms results in non-rectangular boundaries for the market structure outcomes – the $A(D^*, x, z, \theta)$ in equation (25). This leads him to employ a frequency simulation approach to maximum likelihood estimation of the parameters. His estimates suggest strong returns to differentiation. That is, that entry by the same quality rival causes profits to fall much more than if a different quality rival enters. Additionally, he finds that the choice of equilibrium concept appears to have little consequence for his estimates or predictions.

### 3.3.2 Application: Airline City-Pair Entry

Several papers have developed econometric models of airlines’ decisions to serve airline routes, including Reiss and Spiller (1989), Berry (1992) and Ciliberto and Tamer (2003). Unlike Mazzeo (2002), Berry (1992) develops a sequential-move entry model that allows for observed and unobserved firm heterogeneity.

For example, in several specifications Berry estimates a profit function with a homogeneous variable profit function for firm $j$ in market $m$ of

$$V(N_m, X_m) = X_m \beta - \delta \ln(N) + \epsilon_m, \quad (32)$$

\footnote{For instance,}
and a heterogeneous fixed cost term

\[ F_{mj} = Z_{jm} \alpha + \epsilon_{mj}. \] (33)

In these equations, \( X_m \) is a vector that includes distance between the endpoint cities and population, \( \epsilon_{m0} \) is a normally distributed unobserved shock to all firms’ variable profits, the \( Z \) are fixed cost variables that include a dummy for the potential entrant already serving both endpoints, and \( \epsilon_{mj} \) is an independent error in each firm’s profits.\(^{18}\)

A key simplifying assumption in this specification is that only the total number of firms affects profits, meaning that firms are symmetric post-entry. This allows Berry to simplify the calculation of sequential-move equilibria. In his estimations, Berry uses the simulated method of moments approach of (McFadden 1989) and (Pakes and Pollard 1989). At candidate parameter values, he simulates and orders profits of the \( M \) potential entrants

\[ \pi_1 > \pi_2 > \ldots > \pi_M. \] (34)

He then uses the fact that the equilibrium number of firms, \( N^* \), must satisfy

\[ V(N^*, x, z, \theta) + \epsilon_N \geq 0, \] (35)

and

\[ V(N^* + 1, x, z, \theta) + \epsilon_{N^*+1} \leq 0. \] (36)

Because of his symmetric competitor and variable profit assumptions, Berry can guarantee that there will be a unique \( N^* \) for any given set of profit parameters. An issue he does not address is what other types of profit specifications would guarantee a unique \( D^* \) equilibrium. Reiss (1996) considers a case where uniqueness of equilibrium is guaranteed by an assumption on the order of moves. In this case, full maximum likelihood estimation may be possible.

Besides modeling the equilibrium number of firms in the market, Berry’s approach could be used to model the decisions of the individual potential entrants. To do this, one would simulate unobserved profit draws from \( \phi(\epsilon, x, z, \theta) \) and then construct the probability of entry \( \bar{D}_j \) for each potential entrant. In practice this would mean

\(^{18}\)Although Berry motivates the endpoint variable as affecting fixed costs, there are a number of reasons why the scale of operations could also affect variable profits. For example, airline hubs might allow airlines to pool passengers with different destinations on the same planes, letting the airlines use larger, more efficient aircraft and letting them offer more frequent flights.
estimating via a frequency or smooth simulator

\[ D_j(x, z, \theta) = \int_{\cup A_j} \phi(\epsilon) d\epsilon, \quad (37) \]

where \( \cup A_j \) are all regions where firm \( j \) enters. The observed differences between the firms’ decisions and the model’s (simulated) predictions

\[ \bar{D}_j - \bar{D}_j(x, z, \theta) = \nu_j \quad (38) \]

can then be used to form moment conditions.

Table 7 reports some of Berry’s parameter estimates.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Ordered Probit</th>
<th>Firm Probit</th>
<th>Full Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>1.0</td>
<td>-3.4</td>
<td>-5.3</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.35)</td>
</tr>
<tr>
<td>Population</td>
<td>4.3</td>
<td>1.2</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>(0.1)</td>
<td>(0.08)</td>
<td>(0.24)</td>
</tr>
<tr>
<td>Distance</td>
<td>-0.18</td>
<td>1.2</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.17)</td>
<td>(0.3)</td>
</tr>
<tr>
<td>Serving Two Endpoints</td>
<td>–</td>
<td>2.1</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.30)</td>
<td></td>
</tr>
<tr>
<td>Endpoint Size</td>
<td>–</td>
<td>5.5</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.45)</td>
<td></td>
</tr>
<tr>
<td>( \delta )</td>
<td>1.8</td>
<td>–</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.12)</td>
<td></td>
</tr>
</tbody>
</table>

The results suggest that both firm heterogeneity and competition are important determinants of firm profit, so that simple models like bivariate probit (without competitive effects) and ordered probit (without firm heterogeneity) are not adequate descriptions of the industry.

The differences between the full model and the simpler variants of the model raise the question of what fundamental source of variation in the data allows us
to possibly separately estimate the effects of firm heterogeneity and competition. In this particular application, Berry (1992) suggests the importance of variation in the number of potential entrants. In the simple ordered probit, this has no effect other than raising the maximum number of possibly observed firms. In the multivariate probit style model without competitive effects, the probabilities of various observed numbers of firms are generated by the properties of order statistics, which vary in precise ways as the number of potential firms varies. The model with both heterogeneity and competition allows for more realistic entry patterns. Berry and Tamer (2006) pursue this line of reasoning more formally.

In concluding this subsection, we should emphasize that these examples illustrate only some of the potential compromises a researcher may have to rule out multiple equilibria. We should also emphasize that eliminating multiple equilibria in econometric models should not be an end in and of itself. For example, simply assuming firms move in a particular order may result in inconsistent parameter estimates if that order is incorrect. Moreover, it may well be that the multiplicity of pure-strategy equilibria are a fact of life and something that the econometric model should allow. In the next section we shall illustrate approaches that allow multiple outcomes.  \[19\]

### 3.4 Imperfect Information Models

So far we have considered entry models in which potential entrants have perfect information about each other and the researcher has imperfect information about potential entrants’ profits. In these models, the presence of multiple or no pure-strategy equilibria can pose nontrivial identification and estimation issues.

A natural extension of these models is to assume that, like the econometrician, potential entrants have imperfect information about each others’ profits. In this case, potential entrants must base their entry decisions on expected profits, where their expectations are taken with respect to the imperfect information they have about competitors’ profits. As we show below, sometimes the introduction of expectations about other players’ profits may or may not ameliorate multiplicity and non-existence problems.

\[19\]See also Bresnahan and Reiss (1991a) for an analysis of a game with mixed strategies.
3.4.1 A Two-Potential Entrant Model

To appreciate some of the issues that arise in imperfect information models, it is useful to start with a 2 × 2 entry game format of Bresnahan and Reiss’ perfect information model. Following Bresnahan and Reiss’ notation, assume that the heterogeneity in potential entrants’ profits comes in fixed costs. The two firms’ ex post profits as function of their competitor’s entry decision, $D_j$, can be represented as:

$$\pi_1(D_1, D_2) = D_1(\pi^M_1 + D_2 \Delta_1 - \epsilon_1)$$

$$\pi_2(D_1, D_2) = D_2(\pi^M_2 + D_1 \Delta_2 - \epsilon_2),$$

where the $\Delta_i$ represent the effect of competitor entry. To introduce private information in the model, imagine that firm $i$ knows its own fixed cost unobservable $\epsilon_i$ but that it does not know its competitor’s fixed cost unobservable $\epsilon_j$. Assume also that firm $i$ has a distribution $F_i(\epsilon_j)$ of beliefs about the other player’s unobservable fixed costs $\epsilon_j$.

Following the perfect information case, we must map the potential entrants’ latent profit functions into equilibrium strategies for each potential entrant. Unlike the perfect information case, the potential entrants maximize expected profits, where they treat their competitor’s action $D_j$ as a function of the competitor’s unknown fixed costs. Mathematically, firms 1 and 2 enter when their expected profits are positive, or

$$D_1 = 1 \iff D_1(\pi^M_1 + p^1_2 \Delta_1 - \epsilon_1) > 0$$

$$D_2 = 1 \iff D_2(\pi^M_2 + p^2_1 \Delta_2 - \epsilon_2) > 0$$

and $p^j_i = E_i(D_j)$ denotes firm $i$’s expectation about the probability firm $j$ will enter. In equilibrium, these probabilities must be consistent with behavior, requiring

$$p^2_1 = F_1(\pi^M_2 + p^1_2 \Delta_2)$$

$$p^1_2 = F_2(\pi^M_1 + p^2_1 \Delta_1).$$

To complete the econometric model, the researcher must relate their information to the potential entrants’ information. Absent application-specific details, there are many possible assumptions that can be entertained.

One leading case is to assume that the researcher’s uncertainty corresponds to the firms’ uncertainty. In this case, the econometric model will consist of the inequalities (40) and the probability equalities (41).

Because the equations in (41) are nonlinear, it is not immediately straightforward to show that the system has a solution or a unique solution. To illustrate the
nonuniqueness problem, suppose the firms’ private information has a $N(0, \sigma^2)$ distribution, $\pi_1^M = \pi_2^M = 1$ and $\Delta_1 = \Delta_2 = -4$. In the perfect information case where $\sigma^2 = 0$, this game has two pure strategy Nash equilibria ($\{D_1 = 1, D_2 = 0\}$ and $\{D_1 = 0, D_2 = 1\}$) and one mixed strategy equilibrium (both firms enter with probability 0.25). When $\sigma^2$ is greater than zero and small, there is a unique symmetric equilibrium where each firm enters with a probability slightly above 0.25. There also are two asymmetric equilibria, each with one firm entering with a probability close to one and the other with a probability close to zero. These equilibria parallel the three Nash equilibria in the perfect information case. As $\sigma^2$ increases above 1, the asymmetric equilibria eventually vanish and a unique symmetric equilibrium remains. This equilibrium has probabilities $p_1 = p_2$ approaching 0.5 from below as $\sigma^2$ tends to infinity.

Thus, in this example there are multiple equilibria for small amounts of asymmetric information. The multiple equilibria appear because the model’s parameters essentially locate us in the center rectangle of Figure 1, apart from the firm’s private information. If, on the other hand, we had chosen $\pi_1^M = \pi_2^M = 1$ and $\Delta_1 = \Delta_2 = -0.5$ then we would obtain a single symmetric equilibrium with probability $p_2 = p_1^2$ approaching 0.5 from above as $\sigma^2$ tends to infinity and both players entering with probability 1 as $\sigma^2$ tends to zero. (The later result simply reflects that duopoly profits are positive for both firms.) These examples illustrate that introducing private information does not necessarily eliminate the problems found in complete information games. Moreover, in these games there need be no partition of the error space that uniquely describes the number of firms. Finally, any uncertainty the econometrician has about the potential entrants’ profits above and beyond that of the firms will only tend to compound these problems.

### 3.4.2 Quantal Response Equilibria

The above model extends the basic Bresnahan and Reiss duopoly econometric model to the case where potential entrants have private information about their payoffs and the econometrician is symmetrically uninformed about the potential entrants’ payoffs. Independently, theorists and experimentalists have developed game-theoretic models to explain why players might not play Nash equilibrium strategies in normal form games. The quantal response model of McKelvey and Palfrey (1995) is one

\[\text{In the above example, uniqueness appears to be obtainable if the researcher focuses on symmetric equilibria or if restrictions are placed on } \Delta.\]
such model, and it closely parallels the above model. The motivation offered for the quantal response model is, however, different.

In McKelvey and Palfrey (1995), players’ utilities have the form

\[ u_{ij} = u_i(D_{ij}, p_{-i}) + \epsilon_{ij} \]

where \( D_{ij} \) is a strategy of player \( i \) and \( p_{-i} \) represents the probabilities that player \( i \) assigns to the other players playing each of their discrete strategies. The additive strategy-specific error term \( \epsilon_{ij} \) is described as representing “mistakes” or “errors” the agent makes in evaluating the utility \( u_i(\cdot, \cdot) \). The utility specification and its possibly nonlinear dependence on the other players’ strategies is taken as a primitive and is not derived from any underlying assumptions about preferences and player uncertainties.

A quantal response equilibrium (QRE) is defined as a rational expectations equilibrium in which the probabilities \( p_{-i} \) each player assigns to the other players’ playing their strategies is consistent with the probability the players’ play those strategies. Thus, the probability \( p_{ij} = \Pr(u_{ij} \geq \max_k u_{ik}) \) for \( k \neq j \) must match the probability that other players assign to player \( i \) playing strategy \( D_{ij} \).

This quantal response model is very similar to the private information entry model in the previous section. The two are essentially the same when the utility (profit) function \( u_i(D_{ij}, p_{-i}) \) can be interpreted as expected utility (profit). This places restrictions the way the other players’ strategies \( D_{-i} \) enter utility. In the duopoly entry model, the competitor’s strategy entered linearly, so that for example \( \mathbb{E}u_1(D_1, D_2) = u_1(D_1, ED_2) = u_1(D_1, p_2^j) \). In a three-player model, profits (utility) of firm 1 might have the general form

\[ \pi_1(D_1, D_2, D_3) = D_1(\pi_1^M + D_2\Delta_{12} + D_3\Delta_{13} + D_2D_3\Delta_{123}). \]

In this case, independence of the players’ uncertainties (which appears to be maintained in quantal response models) would deliver an expected profit function that depends on the player’s own strategy and the \( p_j^i \).

3.4.3 Asymmetric Entry/Location Models

Quantal response models have been used to model data from a variety of experiments in which players have discrete strategies (see for example Goeree and Holt (2000)). As in McKelvey and Palfrey (1995) empirical work, interest often centers on estimating the variance of the errors in utility as opposed to parameters of the utility functions,
The estimated variance sometimes is used to describe how close the quantal response equilibrium is to a Nash equilibrium. A standard modeling assumption is that the utility errors have a Type 1 extreme value distribution and thus that the choice (strategy) probabilities can be calculated using a scaled logistic distribution. Most studies are able to identify the variance of $\epsilon_{ij}$ because they are modeling experimental choices in which the players’ utilities are presumed to be the monetary incentives offered as part of the experiment.

Seim (2000) and Seim (2004) introduces asymmetric information into an econometric models of potential entrants’ location decisions. Specifically, Seim models a set of $N$ potential entrants deciding in which one, if any, of $L$ locations they will locate. In Seim’s application, the potential entrants are video rental stores and the locations are Census tracts within a town.

In Seim’s model, if potential entrant $i$ enters location $l$, it earns

$$\pi_{il}(\bar{n}_l, x_l) = x_l\beta + \theta_{il} \sum_{j \neq i}^N D_{jl} + \sum_{h \neq l} \theta_{lh} \sum_{k \neq i} D_{kh} + \nu_{il}$$

(42)

where $D_{kh}$ denotes an indicator for whether store $k$ has chosen to enter location $h$; $\bar{n} = n_0, \ldots, n_L$ denotes the number of competitors in each location (i.e., $n_h = \sum_{j \neq i} D_{jh}$); location 0 is treated as the “Do not enter any location.”; $x_i$ is a vector of profit shifters for location $l$; and $\beta$ and $\theta$ are parameters. The own-location effect of competition on profits is measured by the parameter $\theta_{ll}$, while cross-location effects are measured by $\theta_{lh}$. The term $\nu_{il}$ is a store/location specific shock that is observed by the store but not by its rivals. Aside from $\nu$, all of the stores’ profits (in the same location) are identical.

Because a given store does not observe the other stores’ $\nu$’s, the store treats the other stores’ $D_{jh}$ as random variables when computing the expected number of rival stores in each location $h$. By symmetry, each store’s expectation that one of its rivals will enter location $h$ is $p_h = E(D_{kh})$. Given $N - 1$ rivals, the number of expected rivals in location $h$ is then $(N - 1)p_h$.

The assumed linearity of $\pi$ in the $D_{kh}$ is especially convenient for taking the expectation of profits with respect to the equilibrium distribution of rivals. Specifically,

---

21The econometrics of such models are considered further in Aradillas-Lopez (2005).

22By way of comparison, the Bresnahan and Reiss model assumes there is only one location in town.
expected profits for firm $i$ equal

$$E_D\pi_{il}(\bar{n}, x_i) = x_i\beta + \theta_{il}(N - 1)p_l + \sum_{j \neq l} \theta_{lj}(N - 1)p_j + \nu_{il} = \bar{\pi}_l + \nu_{il}. \quad (43)$$

For simplicity, one could assume that the $\nu$'s have the type 1 extreme value or “double-exponential” distribution that leads to multinomial logit choice probabilities. In this case, equation (43) defines a classic logit discrete choice problem. The $L + 1$ entry probabilities, the $p_l$, then map into themselves in equilibrium. These entry probabilities then appear in the stores’ profit and best response functions.

To calculate the Nash equilibrium entry probabilities $p_1, \ldots, p_L$ we must solve the nonlinear system of equations:

$$p_0 = \frac{1}{1 + \sum_{l=1}^L \exp(\bar{\pi}_l)}$$
$$p_1 = \frac{\exp(\bar{\pi}_1)}{1 + \sum_{l=1}^L \exp(\bar{\pi}_l)}$$
$$: \quad : \quad :$$
$$p_L = \frac{\exp(\bar{\pi}_L)}{1 + \sum_{l=1}^L \exp(\bar{\pi}_l)} \quad (44)$$

Seim argues that for fixed $N$, an solution to this equilibrium system exists and is typically unique, although as the relative variance of the $\nu$’s declines, the problem approaches the discrete problem and it seems that the non-uniqueness problem faced in perfect information simultaneous-move could reoccur. The single location model discussion above illustrates how and why this could happen.

Three other noteworthy issues arise when she estimates the model. The first is that in principle she would like to parameterize the scale of the logit error so that she could compare the model estimates to a case where the potential entrants had perfect information about each other’s profits. Unfortunately, this cannot be done because the scale parameter is not separately identified from the profit parameters. The second issue is that the number of potential entrants in each market is unknown.

$^{23}$In practice, Seim treats “no entry” as location 0 in the choice problem and uses a nested logit model of the unobservables, where the entry locations $l > 0$ are more “similar” than the no-entry location.

$^{24}$Seim’s expressions differ somewhat because she uses a nested logit and includes market unobservables.
Seim deals with this problem by making alternative assumptions about the number of potential entrants. The third is that some markets have small census tract locations and others large tracts. To be able to compare $\theta_{ij}$'s estimates across markets and to reduce the number she has to estimate, Seim associates the $\theta$'s with distance bands about any given location. Thus, two neighboring tracts, each ten miles away, would have their entrants weighted equally in a store's profit function. Naturally, competitors in nearer bands are thought to have greater (negative) effects on store profits than competitors in more distant bands, however, these effects are not directly linked to the geographic dispersion of consumers (as for example in the retail demand model of Davis (1997)).

Turning to estimation, the system (44) produces probabilities of entry for each firm for each location in a market. The joint distribution of the number of entrants in each location $\bar{n} = n_0, \ldots, n_L$ is given by the multinomial distribution

$$P(n_0, \ldots, n_L) = N! \prod_{j=0}^{L} p_j^{n_j} n_j!.$$  

Because the probabilities $p_j$ depend on each other, some type of nested fixed point algorithm will have to be used to evaluate the likelihood function for each new parameter vector. Similarly, generalized method of moment techniques that used the expected number of entrants in each location, combined with a nested fixed point algorithm, could be used to compute parameter estimates.

### 3.5 Entry in Auctions

The IO literature has recently devoted considerable attention to estimating structural econometric models of auction participants' bids. Almost all of these empirical models presume that the number of participants is exogenously given. The number of bidders in an auction is then used as a source of variation to identify parameters of the auction model. Recently, there have been several attempts to develop and estimate structural models of auction participation decisions. These papers build on theoretical models that explore how participation and bids are related (e.g., McAfee and McMillan (1987)), Levin and Smith (1994), and Pevnitskaya (2004)).

Auctions are a natural pace in which to apply and extend private information models of entry. This is because the actual number of bidders is typically less than the total eligible to bid in any given auction. Additionally, in the standard auction set-up, bidders are presumed to have private information about their valuations, which
is analogous to potential entrants having private information about costs. As the theoretical literature has emphasized, auction models differ considerably depending upon the affiliation of participants’ information, auction formats and auction rules. In addition, when considering entry, there is an issue of what players know when they bid versus what they know when they make their entry decisions.

To illustrate parallels and differences between auction participation and market entry models, consider a sealed-bid, private values auction with symmetric, risk-neutral bidders. In a first stage, assume that \( N \) potential bidders decide whether to pay a known entry cost \( K \) to enter, and in a second stage, they bid after learning the number of “entering” bidders, \( n \).

Conditional on entering and having \( n - 1 \) rivals, bidder \( i \) with private value \( v_i \) maximizes expected profits of

\[
\pi_i(v_i, b, n) = (v_i - b)\Pi_{j \neq i}^n G(b, j)
\]

by choosing \( b \). In this expression, \( G(b, j) \) is the probability that bidder \( j \) bids less than \( b \). The first-order conditions for this maximization, along with any boundary conditions, determine the optimal bid functions as a function of the private information \( v_i \) and the number of bidders, \( n \). Inserting these bid functions back into the profit function (3.5), delivers an expected equilibrium profit function \( \pi(v_i, n) \) as a function of \( n \).

To predict how many firms \( n \) will bid, we now need to know the timing of the private information. In Levin and Smith, for example, the \( N \) potential bidders do not know their valuations before they sink an entry cost \( K \). In a symmetric equilibrium, the potential bidders randomize their entry decisions wherein they decide to pay the entry cost \( K \) with probability \( p^* \). In equilibrium then, expected profits

\[
\sum_{n=1}^{N} Pr(n - 1, N - 1)E_v\pi(v_i, n) - K
\]

will equal zero. Here, \( E_v \) denotes the expectation with respect to the (symmetric) distribution of private values revealed post-entry. The term \( Pr(n - 1, N - 1) \) denotes the probability that \( n - 1 \) of the remaining \( N - 1 \) firms choose to enter. In a symmetric equilibrium, this probability is the binomial probability

\[
Pr(n - 1, N - 1) = \binom{N - 1}{n - 1}p^{n-1}(1 - p)^{N-n}.
\]
This probability can serve as the basis for estimating \( p^* \) from the empirical distribution distribution of \( n \). In turn, estimates of \( p^* \) can be used to obtain an estimate of \( K \) \(^{25}\).

### 3.6 Other Kinds of Firm Heterogeneity

The empirical applications we have discussed either leave the firm heterogeneity as part of fixed cost, or else model potential heterogeneity as a discrete set of firm types. There are a wide range of choices about endogenous market structure that we have not considered here but are obviously empirically important and would be very useful extensions to the existing empirical literature on “structural” models of market structure. These extensions would include

- endogenous scale of operations,
- endogenous product characteristics in a continuous space,
- endogenous product product quality

Each of these topics is discussed at great length in the theoretical literature in IO and each of these topics has featured in descriptive empirical work on actual industries, but the tie between the theory and empirical work is yet to be well-established.

### 3.7 Dynamics

In a separate chapter in this Handbook, Ariel Pakes discusses a class of dynamic industry models that are intended to be empirically applied. As in Ericson and Pakes (1995), these models incorporate firm heterogeneity, endogenous investment (with uncertain outcomes), imperfect competition and entry and exit. However, to date the empirical application of those models has been somewhat limited and the present empirical applications rely on calibration as much or more than estimation (e.g. Benkard (2004).) One reason for this is the “curse of dimensionality” that makes computation take a long time. There are two solutions to this problem. The first is the development of faster computational techniques and faster computers. The second is the development of econometric techniques to estimate some parameters without fully solving the dynamic model. In the context of the simpler entry models

\(^{25}\)Other estimation strategies are possible.
of this chapter, that is like estimating some demand and marginal cost parameters from the second-stage competition game, without having to solve the entry (but likely using instrumental variables to control for endogenous market structure.)

There are a range of possible models that fall between the strictly static (or “cross-sectional”) models of this chapter and the more complicated dynamic models exemplified by Ericson and Pakes (1995). A first set of steps in this direction is taken by Aguirregabiria and Mira (2004) and by Pakes, Ostrovsky, and Berry (2004) and Pesendorfer and Schmidt-Dengler (2004). For example, one could consider a repeated entry game, where the variable profit function is symmetric (as in Bresnahan and Reiss) and heterogeneity only enters fixed cost (as in Berry (1992).) One might assume that a fraction of the fixed costs are sunk and some fraction must be paid anew each period. To keep the model simple, one might follow Seim (2004) in assuming that each period’s unobservable is i.i.d. and privately observed by the firm. In the dynamic context, the resulting \textit{ex post} regret will affect future entry and exit decisions. One could track data on the number of firms in each market in each time period, learning about sunk costs via the degree to which history matters in predicting \( N \) as a function of current market conditions and past \( N \). Such a model would be much easier to compute than models with richer notions of firm heterogeneity and investment, but of course this would come at the cost of considerable realism. By a series of such steps the literature of this chapter might advance toward the realism of the more complicated dynamic models, without sacrificing a more immediate empirical pay-off.

4 Conclusion

The models of this chapter use the logic of revealed preference to uncover parameters of profit functions from the cross-sectional distribution of market structure (i.e. “entry decisions”) of oligopolist firms across markets of different sizes and types. We have highlighted the role that assumptions on functional form, distributions of unobservables and the nature of competition play in allowing us to estimate the parameters of underlying profits. Considerable progress has been made in applying these models to an increasingly rich set of data and questions. Considerable work remains in dealing with important limitations of the work, including difficult questions about dynamics and multiple equilibria.
References


Figure 1. Monopoly and Duopoly Entry Thresholds

\[ \pi_2^M \]

- No Entrants
- Firm 1 Monopoly
- Firm 2 Monopoly
- Duopoly

\[ \pi_1^M \]
Figure 2. Multi-Stage Sequential Monopoly Model

\[ \pi_2^M = \pi_1^M \]

- No Entrants
- Firm 1 Monopoly
- Firm 2 Monopoly
- Duopoly

\[ 0 \]

\[ \pi_1^M \]

\[ \pi_2^M \]