Empirical Models of Imperfect Competition: A Discussion

By

Liran Einav
Stanford University and NBER

and

Aviv Nevo
Northwestern University and NBER

March 7, 2006

* We wish to thank Ignacio Esponda, Igal Hendel, and Jon Levin for comments. Einav gratefully acknowledges financial support from the National Science Foundation and the hospitality of the Hoover Institution. Nevo gratefully acknowledges financial support from the National Science Foundation and the Sloan Foundation.
Empirical Models of Imperfect Competition: A Discussion

Liran Einav
Stanford University and NBER
leinav@stanford.edu

Aviv Nevo
Northwestern University and NBER
nevo@northwestern.edu

March 7, 2006

Keywords: Auction, entry, price competition, estimation, empirical models

JEL classification: C1, C7, D8, L0

1 Introduction

The field of Industrial Organization (IO) studies the behavior of firms and the interaction among them. In the last 25 years, IO studies have increasingly focused on single industries, using a combination of economic theory and statistics to analyze strategic interaction between firms. The focus on a particular industry allows the researcher to develop a model that takes into account the specific details of the industry. IO economists then use the model to derive comparative statics and test them in the data, and often estimate the structural parameters of the model using state-of-the-art econometric methods.

IO studies a broad array of decisions made by firms, starting from long-run decisions, such as those of entry into particular markets or those that relate to product design and development, medium-run decisions, such as contractual relationships and production, and short-run decisions, such as pricing and bidding in auctions. As each of these decisions is somewhat distinct from others in various aspects (the nature of the decision, the relevant policy questions, the typical data sets available to the empirical researcher, and so on), the literature of recent years can be classified according to which decision it analyzes.

Two issues of particular interest have been entry and exit decisions that determine market structure and price competition. These issues are the focus of the two excellent surveys by Athey and Haile, and Berry and Tamer. Athey and Haile survey the key principles guiding the studies of auction markets, while the paper by Berry and Tamer addresses recent work and identification in empirical models of strategic entry. While the papers and the topics are quite different, they share several themes. First, they exploit the assumption that agents are acting optimally (i.e. maximizing profits or utility) and that the data is generated by equilibrium behavior in order to infer unobserved quantities from observed variables. Second, by building complete economic

---

*We wish to thank Ignacio Esponda, Igal Hendel, and Jon Levin for comments. Einav gratefully acknowledges financial support from the National Science Foundation and the hospitality of the Hoover Institution. Nevo gratefully acknowledges financial support from the National Science Foundation and the Sloan Foundation.
models and estimating their parameters, the work surveyed here permits counterfactual policy experiments. For example, in studies of auctions, data on bids are used to infer unobserved valuations, which can be used to assess, say, counterfactual revenues from alternative auction formats. Similarly, in entry models, observed entry behavior is used to back out fixed costs and intensity of competition, which can then be used to assess the change in market structure in response to various government interventions. Third, both papers emphasize the importance of non-parametric identification by addressing the question of how much can be learned from data without specific functional form assumptions. Both papers suggest that even if eventually the model is estimated parametrically, non-parametric identification is important; it provides the researcher the assurances that with enough data the parametric assumptions required for estimation could be relaxed.

The papers contain a vast amount of material in them, making it impossible to comment in detail on every section. Therefore, our plan in this discussion is to re-iterate some of the key principles the papers emphasize, and to try to place them within a greater context of the field. We hope to achieve two goals in this discussion. The first goal (Section 2) is to survey the main issues and developments in some of the main areas of study, primarily targeting readers who have not followed the field carefully. We briefly summarize both auctions and entry literatures, as well as the demand/pricing literature; the latter is useful as background for our subsequent discussion.

The second goal (Section 3) is to offer a conceptual framework that encompasses the different aspects of firms’ behavior. In particular, we emphasize a conceptual distinction between the way modeling of auction markets (and, to a similar extent, entry) evolved, compared to the empirical models used to study price competition in differentiated product oligopoly. A simple conceptual framework may be appealing for readers outside of IO, who care less about the details of particular IO problems, but want to learn more generally about the key features of empirical IO. In addition, we think that active IO researchers may also benefit from thinking along the lines we sketch in Section 3; to the extent that different areas of research within IO evolve somewhat separately, it is useful to step back and ask whether there is a scope for “arbitrage.”

2 Summary of several areas of study in IO

2.1 Empirical models of price setting

A large literature in empirical IO over the past 25 years has studied short term, price or quantity, competition in a variety of industries (Porter, 1983; Bresnahan, 1987; Berry, Levinsohn, and Pakes, 1995; Nevo, 2001; and many others). The main goal of this literature is to study the form of competition, understand firm behavior, generate a counterfactual, such as the likely effect of a merger, or quantify welfare gains from, say, the introduction of a new product. Although this literature was not discussed in either of the papers in the session, we describe it as background for the discussion in Section 3.

The typical data set will include a cross section, time series, or panel of markets. For each market the researcher observes the quantity sold, the price charged, and possibly advertising for each product. In addition, both market and product characteristics are observed, and in some
cases consumer level data are available.

The main challenge is to infer unobserved marginal costs. The solution is to use quantity and price data to recover demand, and use the demand estimates jointly with the optimality conditions for pricing to back out implied marginal costs. For example, consider a market where single product firms with constant marginal costs set prices to solve

\[
\max_{p_i} (p_i - c_i) D_i(p_i, p_{-i}).
\]

Consequently, in a Nash Equilibrium this choice satisfies the following first order condition:

\[
c_i = p_i + \left( \frac{\partial D_i(p_i, p_{-i})}{\partial p_i} \right)^{-1} D_i(p_i, p_{-i}).
\]

Given estimates of the demand function, \(D_i(p_i, p_{-i})\) and \(\partial D_i(p_i, p_{-i})/\partial p_i\)^{-1} can be computed and used to back out the implied marginal costs. These costs can then be used to test the supply model, as inputs into simulation of a counterfactual, or to fit a marginal cost function.

One important issue in this literature is the specification of the demand function, given the large number of products present in most markets, and the need for a flexible demand model. There are several solutions for the dimensionality problem in the literature. Multi-level demand systems (Hausman, Leonard and Zona, 1994; Hausman, 1997) solve the problem by a-priori separating the products into segments and allowing for flexible functional forms within and across segments. The functional forms can be flexible because the number of products within a segment and the number of segments are relatively small. Discrete choice models (McFadden, 1973; Berry, Levinsohn, and Pakes, 1995) provide an alternative that solves the dimensionality problem by projecting products onto a characteristics space. Thus, the relevant dimension becomes the number of characteristics and not the number of products. Several discrete choice models that allow for flexible substitution patterns across products have been suggested in the literature.

A second important issue is the endogeneity of prices, and potentially other product characteristics, in the estimation of demand. Unless accounted for, this endogeneity may bias the estimated parameters due to possible correlation between prices and unobserved product attributes. The issue arises because firms are assumed to have more information than the researcher. To address this potential problem researchers have formulated the model so that the econometric error term, usually unobserved product characteristics, enters the estimated equation linearly. This makes standard instrumental variable techniques readily applicable. Common instrumental variables include the characteristics of products, which proxy for the degree of competition (Bresnahan, 1981; Berry, Levinsohn, and Pakes, 1995), or prices in other markets, which are correlated through common shocks to marginal costs, and are valid due to an independence assumption on the demand error term (Hausman, 1997; Nevo, 2001).

Recent developments in this area include alternative models of demand, estimation using a combination of individual-level and market-level data (Berry, Levinsohn, and Pakes, 2004), and estimation of dynamic demand (Hendel and Nevo, 2005). Applied work has focused on a variety of questions, including estimation of market power and testing for collusion, simulation of the effects of mergers on prices, regulations and trade constraints, valuation of new goods, computation of price indices, quantifying network effects, vertical relations, and many more.
2.2 Empirical studies of auctions

The main goal of the auction literature surveyed by Athey and Haile is to recover the distribution of bidder valuation (or costs) in order to study the form of competition, understand bidding behavior, and compute the optimal auction design (e.g., the format of the auction or the optimal reserve price). Typical data come from a sequence of similar auctions and include the winning bids and possibly the number of bids, all bids, and/or characteristics of bidders and auctions. A main problem is to infer the unobserved valuations or costs from the observed bids. The solution is to use the known rules of the auction, the optimality of bidding behavior, and the estimated probability of winning in order to back out the unobserved valuations.

Athey and Haile discuss three key ideas in this literature. First, when signals are independent the distribution of a single order static determines the parent distribution. Suppose a researcher observes the winning bid and the number of bidders in an ascending independent private values auctions. Assuming optimal bidding, and that the researcher observes several auctions with the same distribution of valuations, then the distribution of the order statistic can be observed and the distribution of valuations can be recovered non-parametrically.

Second, the parameters of the model can be estimated without the need to solve for the equilibrium bidding strategies, which is often difficult. One possible strategy for estimation would be to compute, for given parameter values, the equilibrium bidding behavior and then choose the parameter values that minimize the distance between observed and predicted behavior. Alternatively, one can avoid computing equilibrium behavior and instead recover it from the data. For example, in a first-price sealed-bid auction with independent private values (IPV) bidders place bids to solve

\[
\max_{b_i} (v_i - b_i) \Pr(b_j \leq b_i \forall j \neq i).
\]  

Consequently, in equilibrium this choice satisfies the following first order condition:

\[
v_i = b_i + \left( \frac{\partial \Pr(b_j \leq b_i \forall j \neq i)}{\partial b_i} \right)^{-1} \Pr(b_j \leq b_i \forall j \neq i).
\]  

The key insight of Guerre, Perrigne, and Vuong (2000) is that if the econometrician observes all the information available to bidders about their competitors, the markdown factor, \( \frac{\partial \Pr(b_j \leq b_i \forall j \neq i)}{\partial b_i} \) \( \Pr(b_j \leq b_i \forall j \neq i) \), can be estimated from observed bids in the same and/or other similar auctions. The observed bids and the first order condition in equation (4) can then be used to recover the unobserved valuation. The recovered valuations can in turn be used to estimate the distribution of valuations. \(^1\)

The third principle discussed by Athey and Haile is the value of additional information. This can be in the form of bidder or auction covariates, or information regarding ex-post valuation.

Athey and Haile touch briefly on some of the recent developments in this area, including application of set identification (Haile and Tamer, 2003), incorporating unobserved heterogeneity (Krasnokutskaya, 2003), dynamics and the importance of capacity constraints (Jofre-Bonet and

\(^1\) If computing equilibrium bidding strategies is feasible, one could imagine iterating the process in order to gain efficiency.
Pesendorfer, 2003), multi-unit auctions and bundling (Cantillon and Pesendorfer, 2004), and auction participation (Athey, Levin, and Seira, 2004). Applied work has addressed questions such as the optimal reserve price, the detection of collusion, testing between private and common value models, and the value of seller reputation.

2.3 Empirical studies of entry

The paper by Berry and Tamer discusses identification in models of strategic entry. The focus of this literature is on the entry and location decisions of firms. The typical data set includes market characteristics, the number of firms, and potentially firm identities for a cross-section of local markets. The goal is to recover the distribution of fixed costs, and the properties of the variable profit function, if data is not available to estimate the variable profit function directly. Applied questions analyze the determinants of firm (or product) entry and exit, the optimal market structure under different scenarios, the speed at which variable profits decline with the number of firms, and the degree of competition and substitution between different market segments.

The basic idea is to back out profitability from entry and location decisions. Firms enter only if it is more profitable than staying out of the market, and in choosing location they choose the most profitable one. There are several unique issues to this literature that stem from the discreteness of the action space in such settings. First, joint non-parametric identification of the distribution of costs and variable profits is difficult. Berry and Tamer provide an excellent discussion of the identification issue, presenting the relevant results from the literature, and extending them as needed. Second, multiplicity of equilibria introduces additional econometric difficulties. In principle, multiple equilibria could arise in both the auctions and the price-setting games previously discussed. In both these cases, however, the estimation is usually based on first order conditions that hold in all equilibria. With discrete controls, as in the entry literature, multiple equilibria are common but restrictions which hold in all equilibria may not exist without further assumptions. Therefore, it is often the case that there is no unique mapping from the observed and unobserved variables to market outcomes. In other words, the model does not generate a unique prediction that can be used for estimation (for example, to compute the likelihood of the data).

Several solutions to the problem of multiple equilibria have been offered in the literature. First, one could search for outcomes that hold in all equilibria, such as the number of firms rather than their identities (Bresnahan and Reiss, 1990; Berry, 1992). This often requires firms to be symmetric in their (post-entry) competitive effect on rivals. A second alternative is to modify the structure of the game so that a unique (potentially probabilistic) prediction is obtained. This can be done by specifying an equilibrium selection mechanism (Mazzeeo, 2002), by assuming a sequential (rather than simultaneous) order of moves (Berry, 1992), by moving from a complete information game to one of incomplete information (Seim, 2005), or by a combination of both (Einav, 2003). A third solution (Ciliberto and Tamer, 2004; Andrews, Berry, and Jia, 2004) is to leave the model incomplete and focus on set identification.

A recent development in this area has been the explicit consideration of dynamics. Most of the earlier work has focused on reduced form variable profit functions. These reduced form profit functions can be used to recover the distribution of entry costs and to possibly learn some
general features of the entry decision, but they cannot be used to distinguish between fixed and sunk costs, or to simulate counterfactual experiments that might change the shape of the reduced form relation. In addition, the explicit modeling of the variable profit function helps with the identification problems discussed by suggesting a framework that can bring in additional information. For example, the variable profit function can be computed from the estimates of the studies discussed above in Sections 2.1 and 2.2.

3 A conceptual framework

In this section we provide a unified conceptual framework for all the areas of research mentioned above. Our goal is to highlight the difference in the nature of the econometric error term as an important distinction between these areas. In what follows we first describe the framework, and then discuss how the literatures relate to it.

3.1 The setup

Consider $N$ players interacting with each other by making choices about a continuous control variable. The econometrician observes two quantities for player $i$: $y_i$ is player $i$’s control variable, and $z_i$ is an outcome variable. We introduce two unobserved (by the econometrician) mean-zero stochastic variables, $\eta_i$ and $\xi_i$, which are associated with player $i$. For simplicity, suppose that both $\eta_i$ and $\xi_i$ for all $i$ are known to all players.

Player $i$’s chooses $y_i$ in order to maximize his objective function given knowledge (or beliefs) about his opponents’ choices, namely to solve

$$\max_{y_i} E \left[ \pi_i(y_i, y_{-i}, \eta_i, \xi_i, \xi_{-i}) | \{ \eta_i, \xi_i \}_{i=1}^{N} \right]. \tag{5}$$

Equation (5) provides the key conceptual distinction between the two stochastic elements, $\eta_i$ and $\xi_i$. While both of them affect player $i$’s choices, only the latter directly affects player $i$’s opponents’ payoffs. This is shown in equation (5) because $\xi_{-i}$ enters the equation, but $\eta_{-i}$ does not. Thus, player $i$ does not care about $\eta_{-i}$ directly. He only uses $\eta_{-i}$ to form beliefs about $y_{-i}$. In contrast, $\xi_{-i}$ directly affects player $i$’s payoffs, in addition to its role in helping player $i$ in forming beliefs about $y_{-i}$. Consequently, we will loosely call $\xi_i$ a “strategic error term” while $\eta_i$ a “non-strategic error term.”

To identify both error terms in the same model, we will need the outcome observation $z_i$, which for example may be a direct observation of $\pi_i(y_i, y_{-i}, \eta_i, \xi_i, \xi_{-i})$. The system can then be inverted so that there would be a one-to-one mapping between $\{(y_i, z_i)\}_{i=1}^{N}$ and $\{(\eta_i, \xi_i)\}_{i=1}^{N}$. Note, however, that in order to establish that this mapping exists and is unique, the function $\pi_i(y_i, y_{-i}, \eta_i, \xi_i, \xi_{-i})$ should satisfy certain regularity conditions.

In contrast, suppose now that $z_i$ is not observed. Then, it is clear that the one-dimensional observations cannot identify a two-dimensional error term without further functional form or distributional assumptions. Suppose then that the strategic error term is assumed away. In such a case, equation (5) simplifies to

$$y_i = \arg \max \pi_i(y_i, y_{-i}, \eta_i) \tag{6}$$
which can be easily inverted, for each player separately, as long as \( \pi_i(y_i, y_{-i}, \eta_i) \) is monotone in \( \eta_i \).

**Example**  As an example, consider the empirical studies of pricing, discussed in Section 2.1. Here \( p_i \) is the control variable and \( q_i = D_i(p_i, p_{-i}) \) is the outcome variable. To fix ideas further, consider the example of a simple logit discrete-choice demand model. The utility for consumer \( h \) from product \( i \) is given by \( u_{hi} = \delta_i - \alpha p_i + \varepsilon_{hi} \) where \( \delta_i \) is the average quality of product \( i \), \( p_i \) is its price, and \( \varepsilon_{hi} \) is an idiosyncratic taste preference, distributed type I extreme value, which is i.i.d across consumers and products. The mean utility from the outside good (good 0) is normalized to zero. If the number of consumers in the market is \( M \), this specification gives rise to the well-known logit demand function:

\[
D_i(p_i, p_{-i}) = M \frac{\exp(\delta_i - \alpha p_i)}{1 + \sum_{j \in J} \exp(\delta_j - \alpha p_j)} \tag{7}
\]

This demand function satisfies the restriction that \( \delta_i \) is a sufficient statistic for player \( i \), so all heterogeneity across firms can be summarized by a one-dimensional parameter. One can think of the unobserved firm-specific marginal cost, \( c_i \), as the non-strategic error term in the conceptual model above, and of \( \delta_i \) as the strategic error term. If we only observed prices, we will not be able to determine whether the price of a certain product is higher because of higher marginal costs (high \( c_i \)) or because of higher quality (high \( \delta_i \)). Clearly, separating these two cases is important for any counterfactual analysis. The latter case, because of its strategic implication, is the one that introduces the problem of price endogeneity in empirical pricing models.

The literature solves the indeterminacy problem by exploiting an additional source of data. We typically observe quantities, as well as prices. Quantities identify the \( \delta_i \)'s, and therefore prices can identify the marginal costs, \( c_i \). Without quantity data, however, the system is not identified, unless we know (or make assumptions about) the product qualities, \( \delta_i \)'s. Even with quantity data, in order to identify the \( \delta_i \)'s we need to assume that \( \delta_i \) is one-dimensional sufficient statistic for player \( i \) and that the distribution of \( \varepsilon_{hi} \) is either known or restricted by certain parametric assumptions.

### 3.2 Discussion

There are several parallels between the empirical auctions literature and the empirical studies of price setting games. Both are focused on recovering an unobserved primitive, bidder valuation or marginal costs, from observed behavior, bids or prices. Indeed, even the mathematical structure of the problem is similar, as can be seen from the similarity between the structure of equation (4) and equation (2). The similarity is, of course, not incidental; after all, one way of thinking about IPV auctions is as an incomplete information version of a Bertrand price competition with homogeneous products. Despite this striking similarity in the nature of the problem, the two literatures evolved in very different directions. Much of the auction literature emphasizes non-parametric identification and non-parametric estimation techniques, while the demand literature concentrates on parametric ways to deal with endogeneity of prices.
One common belief is that studies of auctions tend to have better data. While there might be a sense in which this is true, it is not obvious. Normally, a researcher interested in estimating demand will have quantities and all prices, while a researcher studying an auction might only observe the winning bid (price). This makes the question even more relevant: what is it that allows the empirical auction literature to focus on non-parametric methods while the literature studying price and quantity competition is mostly parametric?

Part of the answer has to do with the fact that in auctions the rules of the game are known and in many cases sufficient for estimation. For example, suppose that we wanted to study a pricing game similar to an ascending IPV “button” auction. Namely, suppose that firms sell a homogenous good, get private independent draws of marginal costs, which are drawn from an identical distribution, and compete in prices. The object of interest is the distribution of cost, which can be recovered from the observed prices. So in a setting where one believes this is the right model the methods of auctions can be used.

One objection to the setup offered in the previous paragraph for many industries is that products are differentiated. This alone is not enough to explain the difference between the two literatures. For example, suppose one observes a sequence of markets with differentiated products. In each market the products receive independent identical shocks to marginal costs, and there are no other unobserved shocks. If products are symmetric then the demand for a product, as a function of its price, can be recovered non-parametrically even with a large number of products because only the number of competitors matters (but not their identities). Alternatively, if products are not symmetric but are present in all markets, then the demand for each product can be recovered even with a large number of products because the shocks to marginal costs are i.i.d. across markets. As the setup becomes more complicated, however, because, say, products vary in unobserved (to the econometrician) dimensions across markets and some products are better substitutes than others, or if the shocks to marginal costs are not independent across markets, then non-parametric estimation will not be feasible with reasonably sized data sets. Thus, the use of parametric models for price competition is not driven by the dimensionality per se, but rather by the combination of the dimensionality, symmetry assumptions, and a wedge between the information available to the econometrician and that available to the players.

We believe that the key difference between the price-setting and auction literatures relates to the conceptual framework presented earlier. For simplicity, consider an IPV procurement auction. A quick comparison of equation (3) and equation (1) reveals that one can think of the probability of winning as the demand function. If bidders are ex-ante symmetric (or can be a-priori mapped into a finite set of types), the auction model does not give rise to a strategic error term. Since in such a setting the idiosyncratic shocks are private information, they cannot enter opponents’ considerations, and therefore are similar to the \( \eta_i \)’s in the conceptual framework.

One should note, however, that strategic error terms in auction environments may exist by allowing a richer structure of unobserved heterogeneity. As an example, they may show up when there exist differences among bidders, which are common knowledge to all participants (but not to the econometrician). Common knowledge differences are strategic: they make one bidder’s expectation about his probability of winning, given a bid, be different than those expectation of
other bidders, given the same bid. Cost variation which is \textit{private information} is non-strategic: by construction, it does not enter the opponent’s optimization problem.

There may be various reasons why the auction literature has evolved in this way. First, symmetry and exchangeability assumptions are reasonable approximations in many auction settings, but may be less credible approximations in product markets. Second, the parallel to quantity data (i.e. the probability of winning) is, of course, not observable, so identification of a strategic error-term with typical auction data sets and without more parametric assumptions is impossible. At the same time, the absence of “quantity” data does not allow the researcher to falsify the modeling assumptions, as the model is just identified. It may be interesting to analyze whether auxiliary information about outcome variables (e.g. the ex-post resale value of an object, or the cost of a project), which are emphasized by Athey and Haile, may help in testing the typical empirical models, and in identifying richer models, which allow for strategic error terms.

The entry literature may also fit into the conceptual framework. The entry literature does not typically model a strategic error term. The stochastic term is in the sunk cost of entry, which (conditional on the entry decision) does not affect the profitability of opponents. One can imagine a strategic error term here; namely, it is reasonable to think that firms are more likely to enter a market because of lower entry cost, or because of better productivity. The latter is strategic, as better productivity will make such firms more profitable after entry, and their opponents less. In fact, such an additional (strategic) error term may be identified from post-entry price and quantity data, which can often be available in applications of entry models.

4 Concluding remarks

The empirical IO literature has evolved quite rapidly over the last few years. As the papers in the session demonstrated, much progress has been made on identification and estimation of many different dimensions of firms’ decisions. For example, we have more flexible models of consumer demand, better methods to non-parametrically estimate bidder valuation in auctions, and significant progress has been made on estimating entry and dynamic games. Given these important methodological advances it is time to apply these methods. Through systematic application to different industries we will be able to learn about the economy and about how to even further improve our methods.

References


