Empirical Likelihood Estimation of Asymmetric Information Models with an R&D Application

Andres Aradillas-Lopez*

September 26, 2003

Abstract: An empirical likelihood estimator is proposed for models where agents interact under asymmetric information. The methodology focuses on situations where some variables that were privately known when the choices were made becomes available to the econometrician afterwards. The main feature of the estimator is that payoffs parameters, beliefs and the unknown distribution of these privately known variables are estimated simultaneously under the assumption that observed outcomes are the result of a Bayesian Nash equilibrium. The methodology is applied to a model in which firms simultaneously decide to stay in the industry or to leave and (if they decide to stay) they then decide if they increase their R&D expenditure or not. Estimation results show significant strategic R&D interaction. It also shows that this interaction is more important for small firms than for larger ones.
Keywords: *Empirical likelihood, asymmetric information, Bayesian-Nash equilibrium.*

1 Introduction

In models of economic interaction with game theoretical foundations, each agent's payoff is affected by the choices made by other agents. If all agents' payoffs were public information and all agents behaved rationally, each agent would be able to predict others' actions in a (pure strategy) Nash equilibrium. However, this perfect information environment may be implausible in a number of economic settings: agents may try to deliberately conceal their own payoffs, or it may be prohibitively costly for each agent to fully determine what other agents' payoffs are, especially when interaction takes place

*I would like to thank Professors Guido Imbens, Daniel McFadden, Jim Powell, Michael Jansson and Sofia Berto Villas-Boas for helpful comments.

among a relatively large number of agents.

The presence of asymmetric information implies that agents must construct beliefs about other agents' actions. Although these beliefs may be completely heterogeneous and arbitrary, if we assume that the observed choices are derived from a (bayesian) Nash equilibrium, then these beliefs must "be rational" and satisfy the conditions consistent with such an equilibrium. To illustrate these ideas, consider the following simple example of a 2×2 game in which players simultaneously (i.e, before observing their opponent's choice) must choose between one of two actions: "Fight" or "Don't Fight". Without loss of generality, assume the following payoff matrix:

Fig. 1.- A simple 2×2 game

		PLAYER 2	
		Fight	Don't
PLAYER 1	Fight	$t_1 - \alpha_1\,,\, t_2 - \alpha_2$	$t_1\,,\,0$
	Don't	$0\,,\,t_2$	$0\,,\,0$

Now suppose that α_1 and α_2 are known by both players but that t_1 and t_2 are private information, but it is common knowledge that they are both independent random draws from the same -known by both players- distribution, with cdf given by $\mathcal{P}(t)$.

Let:

$$\begin{aligned} \pi_1 &= \text{Probability that player 1 will choose fight.} \\ \pi_2 &= \text{Probability that player 2 will choose fight.} \\ E_1[\pi_2] &\equiv \text{Player 1's belief that player 2 will choose fight.} \\ E_2[\pi_1] &\equiv \text{Player 2's belief that player 1 will choose fight.} \end{aligned}$$

Now let $E_1[u_1^{\text{Fight}}]$ and $E_2[u_2^{\text{Fight}}]$ be the *expected* payoff from playing fight for players 1 and 2 respectively. Then, due to the linearity of the payoff functions, these expected payoffs are simply given by:

$$E_1[u_1^{\text{Fight}}] = t_1 - E_1[\pi_2]\alpha_1 \quad \text{and} \quad E_2[u_2^{\text{Fight}}] = t_2 - E_2[\pi_1]\alpha_2$$

Because the payoffs from not fighting are normalized to zero in this game, we have that players 1 and 2 will choose fight if and only if $E_1[u_1^{\text{Fight}}] > 0$ and

$E_2[u_2^{\text{Fight}}] > 0$ respectively. Therefore, in a Bayesian Nash equilibrium beliefs must satisfy:

$$E_1[\pi_2] = 1 - \mathcal{P}(E_2[\pi_1]\alpha_1) \quad \text{and} \quad E_2[\pi_1] = 1 - \mathcal{P}(E_1[\pi_2]\alpha_2) \qquad (1)$$

Payoff functions are of course unobservable. Suppose t_1 and t_2 can be expressed as functions of $(\boldsymbol{X}_1, \varepsilon_1)$ and $(\boldsymbol{X}_2, \varepsilon_2)$ respectively. The following assumptions preserve the stochastic and informational assumptions of this simple game:

A1.– $\boldsymbol{X}_1 \in \mathbb{R}^k$ and $\boldsymbol{X}_2 \in \mathbb{R}^k$ are independent draws from the same distribution with (joint) cdf given by $F(\boldsymbol{x})$, and corresponding pdf given by $dF(\boldsymbol{x})$.

A2.– $\varepsilon_1 \in \mathbb{R}$ and $\varepsilon_2 \in \mathbb{R}$ are independent draws from the same distribution with cdf given by $G(\epsilon)$.

A3.– ε_p is independent from $\boldsymbol{X}_p$ for $p \in 1, 2$.

A4.– *At the time the game is played*, the realizations of $(\boldsymbol{X}_1, \varepsilon_1)$ and $(\boldsymbol{X}_2, \varepsilon_2)$ are privately known by players 1 and 2 respectively. This is consistent with the following situations:

  A4.1.– Both players deliberately and effectively conceal the true values of $(\boldsymbol{X}_p, \varepsilon_p)$, $p \in 1, 2$.

  A4.2.– It could be possible for player $p \in 1, 2$ to learn the realization of his opponent's $(\boldsymbol{X}_{-p}, \varepsilon_{-p})$ but it is not profitable to do so.

A5.– Distributions $(F(\boldsymbol{x}), G(\epsilon))$ are known by both players.

Without loss of generality, suppose we parameterize t_1 and t_2 as

$$t_1 = \boldsymbol{\beta}'\boldsymbol{X}_1 - \varepsilon_1, \quad t_2 = \boldsymbol{\beta}'\boldsymbol{X}_2 - \varepsilon_2$$

where the parameter vector $\boldsymbol{\beta}$ is known by both players. Then the Bayesian-Nash equilibrium conditions become

$$E_2[\pi_1] = \int_{\boldsymbol{x}} G\big(\boldsymbol{\beta}'\boldsymbol{X}_1 - E_1[\pi_2]\alpha_1\big)dF(\boldsymbol{x}) \quad \text{and} \quad E_1[\pi_2] = \int_{\boldsymbol{x}} G\big(\boldsymbol{\beta}'\boldsymbol{X}_2 - E_2[\pi_1]\alpha_2\big)dF(\boldsymbol{x}) \quad (1')$$

Now, suppose some time after the game was played by a random sample of M pairs of players, the econometrician has access to the M outcomes and suppose that:

$B1.-$ Assumptions $(A1 - A5)$ were satisfied when the game was played by each of the N pairs of players.

$B2.-$ The realizations of $\{\boldsymbol{X}_{1,i}, \boldsymbol{X}_{2,i}\}_{i=1}^{M}$ are now available to the econometrician.

$B3.-$ The realizations of $\{\varepsilon_{1,i}, \varepsilon_{2,i}\}_{i=1}^{M}$ are not observable by the econometrician.

$B4.-$ The distribution $G(\varepsilon)$ is assumed to be known -up to a finite number of parameters- to the econometrician.

$B5.-$ No particular functional form is assumed for the distribution $F(\boldsymbol{x})$. We will only assume that this distribution *does not depend on any of the payoff parameters, beliefs or the unknown parameters of* $G(\epsilon)$.

The methodology proposed here is aimed at the econometric estimation of models that can be characterized by assumptions $B1 - B5$, but it applies to a much broader class of models than the simple one described above. In particular, it can be applied to models in which all agents can belong to one of a *finite* number of "types", and each player's "type" is public information. Players' types contain some information about their private payoffs. This would be the case for example if in the model presented above there exists a partition of $\mathbb{R}^k$, say $\{\mathcal{X}_1, .., \mathcal{X}_T\}$, where $\mathcal{X}_s \cap \mathcal{X}_t = \emptyset$ for all $s \neq t$ and $\mathcal{X}_1 \cup \cdots \cup \mathcal{X}_T = \mathbb{R}^k$. We say that player p belongs to type τ_t if and only if $\boldsymbol{X}_p \in \mathcal{X}_t$. Then, the methodology presented here could be easily adapted to a model in which all players' types are publicly observed. It can also be used to handle signalling games with a finite number of possible types, and also models based on games played between many opponents simultaneously.

The models of interest for this methodology could be seen as asymmetric information versions of the so-called *interactions-based models*, a term used by Brock and Durlauf (2001) to refer to a "class of economic environments in which the payoff function of a given agent takes as direct arguments the choices of other agents". The focus here is on interactions-based models where an individual can't predict others' choices because payoffs are (at least partially) private information but where some part of this private information becomes available afterwards to the econometrician.

For all possible applications, the proposal is to estimate *simultaneously* the following:

1.- The payoff parameters (α_1,α_2 and $\boldsymbol{\beta}$ in the model described above).

2.- Agents' beliefs ($E_1[\pi_2]$ and $E_2[\pi_1]$ in the model described above).

3.- The unknown parameters from the distribution $G(\epsilon)$ of those variables that are privately observed when the game is played and remain unobservable to the econometrician.

3.- The unknown distribution $dF(\boldsymbol{x})$ of those variables that are privately observed when the game is played, but available afterwards to the econometrician.

Estimation will take place assuming that the observed outcomes are the result of a Bayesian-Nash equilibrium. The link between all of these parameters is given by the corresponding equilibrium restrictions that these beliefs must satisfy (equation (1') in the example presented above). The issues of existence and uniqueness of an equilibrium are crucial and they will be addressed, along with the asymptotic properties of the proposed estimator in section 3 below, where we present the methodology for a particular application of an R&D model.

2 Empirical likelihood: brief overview

2.1 Empirical likelihood: basic problems and some extensions.

Empirical likelihood (EL) was formally introduced by Owen (1988, 1990, 1991). In its simplest form, EL was proposed as a device to construct non-parametric tests and confidence intervals for a mean of a random variable $Z \in \mathbb{R}$ with unknown probability distribution function (pdf). Suppose we have a random sample $\{Z_i\}_{i=1}^N$ and we wish to test if $E[Z] = \mu$. EL would then substitute the unknown pdf with a set of weights $\{p_i\}_{i=1}^N$. The optimal weights would be the solution to the problem

$$\underset{\{p_i\}_{i=1}^N}{\text{Max}} \sum_{i=1}^N \log p_i \quad \text{subject to:} \quad p_i \geq 0 \ , \quad \sum_{i=1}^N p_i = 1 \quad \sum_{i=1}^N p_i Z_i = \mu$$

That is, to maximize the empirical log-likelihood $\sum_{i=1}^N \log p_i$ subject to the weights being a well-behaved pdf, and the data obeying $E[Z] = \mu$ with this pseudo-pdf. Without the constraint $\sum_{i=1}^N p_i Z_i = \mu$, it is easy to show that the uniform weights $p_i = (1/N)\ \forall i$ maximize the empirical log-likelihood. These would be the optimal weights if $\mu = \overline{Z}$ (the sample mean of $\{Z_i\}_{i=1}^N$). Let $\ell(\mu) = \sum_{i=1}^N \log \hat{p}_i$ be the corresponding maximum EL and define the empirical log-likelihood *ratio* $\mathcal{R}(\mu)$ as

$$\mathcal{R}(\mu) = -2 \times \left\{ \ell(\mu) - \sum_{i=1}^N \log(1/N) \right\}$$

where $\{\widehat{p}_i\}_{i=1}^N$ are the optimal EL weights. Now let μ_0 be the true mean of Z. Owen showed that under fairly general conditions $\mathcal{R}(\mu_0) \xrightarrow{d} \chi_1^2$. This implies that hypothesis testing and confidence intervals could be based on the statistic $\mathcal{R}(\mu)$. A α-level confidence interval for example, would be constructed as the set of $\mu \in \mathbb{R}$ such that $\mathcal{R}(\mu) \leq c_\alpha$, where $Pr(\chi_1^2) \leq c_\alpha = 1-\alpha$. Note that if we wanted to estimate μ by maximizing $\ell(\mu)$, we would get $\mu = \overline{Z}$, and the corresponding optimal weights would be the uniform weights $\widehat{p}_i = 1/N$.

EL was also applied to deal with moments other than the mean, and to handle vector-valued random variables, where the weights are estimates of a joint pdf. An important extension was done by Qin and Lawless (1994), who applied EL for general estimating equations: Suppose that for a random variable $\boldsymbol{Z} \in \mathbb{R}^d$ there exist a parameter $\boldsymbol{\theta} \in \mathbb{R}^p$ and a vector valued function $m(\boldsymbol{Z},\boldsymbol{\theta}) \in \mathbb{R}^s$ such that $E[m(\boldsymbol{Z},\boldsymbol{\theta})] = 0$. For a fixed $\boldsymbol{\theta}$, the corresponding EL problem is to solve:

$$\operatorname*{Max}_{\{p_i\}_{i=1}^N} \sum_{i=1}^N \log p_i \quad \text{subject to:} \quad p_i \geq 0\ , \quad \sum_{i=1}^N p_i = 1 \quad \sum_{i=1}^N p_i m(\boldsymbol{Z}_i,\boldsymbol{\theta}) = 0$$

Let $\ell(\boldsymbol{\theta}) = \sum_{i=1}^N \log \hat{p}_i$ be the corresponding maximum EL. Letting $\boldsymbol{\theta}_0$ be the true parameter value, Qin and Lawless then showed that under some regularity conditions

$$\mathcal{R}(\boldsymbol{\theta}_0) = -2 \times \left\{ \ell(\boldsymbol{\theta}_0) - \sum_{i=1}^N \log(1/N) \right\} \xrightarrow{d} \chi_q^2$$

where q is the rank of $\mathrm{Var}(m(\boldsymbol{Z},\boldsymbol{\theta}_0))$. Confidence regions can be built and hypotheses can be tested for $\boldsymbol{\theta}$ using the statistic $\mathcal{R}(\boldsymbol{\theta})$. We can also use EL to estimate $\boldsymbol{\theta}$ by maximizing $\ell(\boldsymbol{\theta})$. If p=s, then $\widehat{\boldsymbol{\theta}}$ is simply given by the solution to $\sum_{i=1}^{N} m(\boldsymbol{z}_i, \widehat{\boldsymbol{\theta}}) = 0$ and the resulting optimal weights are the uniform ones $\widehat{p}_i = 1/N$. The interesting case is when $s > p$. The latter case would be the kind of problem econometricians usually analyze using GMM estimation.

EL was also extended to analyze combinations of parametric and empirical likelihoods. Suppose for example that the conditional distribution of $y \in \mathbb{R}$ given $\boldsymbol{Z} \in \mathbb{R}^k$ is assumed to have a known parametric functional form given by $f(y \mid \boldsymbol{Z}, \boldsymbol{\theta})$, but that the marginal pdf of $\boldsymbol{Z}$ is unknown and denote it by $dF(\boldsymbol{z})$. The joint pdf of $(Y, \boldsymbol{Z})$ would then be given by $f(y \mid \boldsymbol{z}, \boldsymbol{\theta}) dF(\boldsymbol{z})$. Suppose now that we know that $E[\psi(\boldsymbol{Z}, \boldsymbol{\theta})] = 0$ for some function $\psi \in \mathbb{R}^p$. EL would estimate $\boldsymbol{\theta}$ and $\{p_i\}_{i=1}^{N}$ by solving:

$$\underset{\boldsymbol{\theta}, \{p_i\}_{i=1}^{N}}{\text{Max}} \sum_{i=1}^{N} \log f(y_i \mid \boldsymbol{z}_i, \boldsymbol{\theta}) + \sum_{i=1}^{N} \log p_i$$

$$\text{subject to} \quad p_i \geq 0 \;, \quad \sum_{i=1}^{N} p_i = 1 \quad \sum_{i=1}^{N} p_i \psi(\boldsymbol{z}_i, \boldsymbol{\theta}) = 0$$

Qin (1994, 2000) called the combination or parametric and nonparametric likelihoods "Semi-Empirical Likelihood". Parametric and empirical likelihoods have also been combined in other settings, as in Qin (1998) for upgraded mixture models where one sample $\boldsymbol{z}_1, ..., \boldsymbol{z}_n$ is directly observed from a distribution $F(\boldsymbol{z})$ while another sample $\boldsymbol{x}_1, ..., \boldsymbol{x}_m$ have density $\int p(\boldsymbol{x} \mid \boldsymbol{z}) dF(\boldsymbol{z})$ where $p(\boldsymbol{x} \mid \boldsymbol{z})$ is parameterized as $p(\boldsymbol{x} \mid \boldsymbol{z}, \boldsymbol{\theta})$. Parametric and empirical likelihoods have also been combined in Bayesian models: Lazar (2000) analyzed the product of a prior density on the univariate mean and an empirical likelihood for that mean.

The methodology proposed here is a particular case of semi-empirical likelihood estimation.

2.1.1 Empirical Likelihood and GMM

Every GMM problem can also be estimated using EL. Asymptotic equivalence to first order of approximation between GMM and EL has been well documented in a variety of settings (Owen (2001) is the best comprehensive reference). It has been also well established that EL improves on the small sample properties of GMM. However, other closely estimators have also been developed that improve on the small sample properties of GMM: Continuous updating (CUE) -also called "euclidean likelihood" by Owen (2001)- and exponential tilting estimators (ET). All these belong to a class of Generalized Empirical Likelihood (GEL) estimators. They all have the same asymptotic distribution as GMM but different higher order asymptotic properties. The natural question would be: Why use EL among the GEL family?

A growing body of literature has been devoted to exploring the higher order asymptotic properties of EL. The majority of these efforts have been aimed at test statistics. EL has been found to have higher order optimality properties consistently better than GMM and at least as good as continuous updating estimators. Kitamura (2001) proves important large deviations optimality results for empirical likelihood vis à vis GMM: Of 32 simulations performed, EL had greatest power 22 times, while 2-step, 10-step and continuous updating did this 5, 7 and 0 times respectively. He also found that EL's power ranking was best for hypotheses farther from the null.

The most relevant results to the problem we address here is Newey and Smith (2001). They compare the properties of GEL and GMM estimators and find that EL has two advantages: First, they show that its asymptotic bias does *not* grow with the number of moment restrictions, while the bias of the other often grows without bound. Second, they show that the bias corrected EL is asymptotically efficient relative to the other bias corrected estimators.

3 Proposed application: R&D model

The methodology presented above can be adapted to a number of different economic situations. Instead of observing n different outcomes of a game played by n different $k-$tuples of players (as in the example of the previous sections) we may observe a single outcome of a game played by n different players simultaneously. The application presented here corresponds to the latter case.

The 2×2 game described above was used to illustrate the properties of the proposed empirical likelihood estimator. A brief description of an R&D

model with asymmetric information is presented here. It involves many players (instead of only two) and beliefs (each player has more than one opponent now). In this model firms must simultaneously make an R&D decision in an environment of asymmetric information. First, I specify what I mean by "an R&D decision" by defining the particular space of actions for this model.

3.1 The model

First, some notation: For firm i denote: $d_i \equiv$ Firm i's industry and $k_i \equiv$ Firm i's technological category. Note that $d_i \in \{1, .., D\}$ and $k_i \in$ {LT,SS,SL,HT}.

3.1.1 Timing of the Firms' Decisions:

In period t, each firm makes the following decisions sequentially:

t.1.- Decide to remain or leave the industry in period $t+1$.

t.2.- If the firm decides to remain in the industry, it must choose to increase or not its R&D investment in period $t+1$ (relative to period t).

All firms make these decisions simultaneously (i.e, before observing what other firms have optimally chosen to do) and in the context of asymmetric information, which will be described below. First, denote firms' decisions by

$$Y_1(i) = \begin{cases} 1 & \text{If firm } i \text{ decides to stay for period } t+1 \text{ and } \textit{increase} \text{ its R\&D investment.} \\ 0 & \text{otherwise} \end{cases}$$

$$Y_2(i) = \begin{cases} 1 & \text{If firm } i \text{ decides to stay for period } t+1 \text{ and } \textit{decrease} \text{ its R\&D investment.} \\ 0 & \text{otherwise} \end{cases}$$

$$Y_3(i) = \begin{cases} 1 & \text{If firm } i \text{ decides to exit for period } t+1. \\ 0 & \text{otherwise} \end{cases}$$

I explain next how firms are assumed to be affected by each others' choices in this particular model.

3.2 Strategic interaction among firms

Because R&D decisions may have long-run implications for firm's future performance, it seems natural to assume that firms care about the choices made by others. The methodology presented here allows us to test precisely this assumption. Firms interact in many dimensions, but because a firm's relative size in its industry has been consistently cited as a determinant of innovation as well as market structure (exit/entry decisions) in existing models -see below-, the present model will attempt to *analyze how small and large firms interact.* Specifically, the goal is to answer the following questions:

i.- Do small firms care about the R&D and exit decisions made by other small firms? Do they care about the decisions made by large firms?

ii.- Do large firms care about the R&D and exit decisions made by other large firms? Do they care about the decisions made by small firms?

Choice rules will be modelled in such a way that allows us to test separately the influence of other firms' exit and R&D choices on a particular firm's decision exit and R&D decisions respectively.

3.3 Decision rules

The superscripts S and L will be used to denote a small and large firm respectively.

Then, for a small firm let

$$\begin{aligned} u^S_{R\&D} &= \alpha^S_1 \pi^S_{R\&D} + \alpha^S_2 \pi^L_{R\&D} + \boldsymbol{\beta}'_{R\&D}\boldsymbol{X}_{R\&D} - \varepsilon_{R\&D} \\ u^S_{Stay} &= \gamma^S_1 \pi^S_{Exit} + \gamma^S_2 \pi^L_{Exit} + \boldsymbol{\beta}'_{Stay}\boldsymbol{X}_{Stay} - \varepsilon_{Stay} \end{aligned} \tag{2}$$

Similarly, for a large firm let:

$$\begin{aligned} u^L_{R\&D} &= \alpha^L_1 \pi^L_{R\&D} + \alpha^L_2 \pi^S_{R\&D} + \boldsymbol{\beta}'_{R\&D}\boldsymbol{X}_{R\&D} - \varepsilon_{R\&D} \\ u^L_{Stay} &= \gamma^L_1 \pi^L_{Exit} + \gamma^L_2 \pi^S_{Exit} + \boldsymbol{\beta}'_{Stay}\boldsymbol{X}_{Stay} - \varepsilon_{Stay} \end{aligned} \tag{3}$$

Where

$\pi^S_{R\&D}$ =Proportion of the population of *Small* firms that will choose to increase their R&D *given that they have chosen to stay.*

π^S_{Exit} =Proportion of the population of *Small* firms that will choose to exit.

the equivalent definitions for large firms follow for $\pi^L_{R\&D}$ and π^S_{Exit}. The remaining variables $\boldsymbol{X}_{R\&D}$, $\varepsilon_{R\&D}$, $\boldsymbol{X}_{Exit}$ and ε_{Exit} will be described below.

3.3.1 Optimal decision rules

If the population proportions $\pi^S_{R\&D}$, π^S_{Exit}, $\pi^L_{R\&D}$ and π^L_{Exit} were known, then:

i.- A small firm would:

1.- Stay for period $t+1$ iff $u^S_{Stay} > 0$.

2.- Increase its R&D investment (given that it has chosen to stay) iff $u^S_{R\&D} > 0$.

ii.- A large firm would:

1.- Stay for period $t+1$ iff $u^L_{Stay} > 0$.

2.- Increase its R&D investment (given that it has chosen to stay) iff $u^L_{R\&D} > 0$.

Due to the asymmetric information nature of the model, the proportions $\pi^S_{R\&D}$, π^S_{Exit}, $\pi^L_{R\&D}$ and π^L_{Exit} are not public information. Firms will then maximize the expected version of the payoff functions (2) and (3). This shall be carefully detailed below.

3.4 Strategic interaction

3.4.1 Interaction coefficients

As it was mentioned above, the goal of the model is to estimate the influence of other firms' choices on an individual firm's R&D and exit decisions.

- For a small firm, the coefficients α^S_1 and α^S_2 indicate the influence of the populations of small and large firms' R&D decisions respectively [1] on the firm's own R&D choice. Similarly, γ^S_1 and γ^S_2 measure the influence of the population of small and large firms' exit decisions on the firm's own exit decision.

- Parallel descriptions apply to the interaction coefficients of a large firm: α^L_1, α^L_2 and γ^L_1, γ^L_2.

[1] That is, the R&D decisions of those firms which will actually choose to stay in business.

3.4.2 Why would firms interact?

R&D choices: Modern models of firm survival argue that a firm's innovation capabilities determine its chances of surviving in the long run. Returns to R&D investment can't be easily measured in the short run, but lagging behind other firms in terms of R&D efforts could imply an irreversible long-run disadvantage for a particular firm. In this sense, *interaction would be based on long-run consequences.*

Exit decisions: A firm's expected future performance depends -partially- on market structure. The hypothesis that underperforming firms may delay their exit if they believe market structure will change favorably is an interesting one to test. This effect hasn't been directly isolated and tested before.

3.5 Determinants of exit and R&D

3.5.1 Firm's relative size

Economies of scale have long been mentioned as a determinant of market structure. Game-theoretical models of size-expansion as entry deterrent are well known and used in IO. The most comprehensive summary of such models was done by Panzar (1989). Falling behind in terms of relative size in a given industry may lead to a firm's exit depending on the existence and importance of economies of scale.

Firm's size has also been analyzed as a determinant of innovation efforts: Larger firms may be able to rip the benefits of innovation more effectively than smaller ones. For example, Hall and Vopel (1997) argue that firms with a larger market share may have higher innovative efforts. According to the authors, this phenomenon occurs because market valuation of a firm's innovative efforts may be higher for firms with a larger market share. Klepper and Simmons (2000) present a model in which large firms tend to pull ahead due to the existence of advantages to scale in R&D. Nevertheless, there exist (older) models that have proposed an opposite effect of market share and innovation: The relative lack of competition threat from other firms may discourage innovation efforts by large, incumbent firm.

3.5.2 Firm's market value

Financial markets value the knowledge assets of firms. Innovation serves as an instrument to balance industry and technological uncertainties -see Oriani, R. and Sobrero, M. (2003)-. A firm's ability to hedge these risks through its R&D decisions is -at least partially- reflected in the firm's market value. Changes in the latter may serve as a guide to R&D investment. See for example Hall (1999).

Because market value reflects the public's appraisal of a firm's overall performance, it comprises at least partially all qualitative and quantitative factors that determine firm's survival. Thus, it may become a variable that determines if a firm stays in an industry or not.

3.5.3 A firm's ability to innovate

Successful innovation leads to technological change that may affect market structure: Jovanovic and MacDonald (1994) present a model in which a process of innovation (introduced from outside the industry) increases the minimum efficient scale and induces the exit of firms who fail to master the new technology -see also Klepper (1996)-. On the other hand, successful innovation and the knowledge derived from it may also increase the returns to future innovation efforts (R&D expenditure). Reinganum (1989) proposes such a model.

3.5.4 Technological characteristics

The degree and nature of R&D expenditure and inter-firm competition depends on the technological nature of each industry. Chandler (1994) presents an historical account and classifies U.S industries into three technological groups or segments:

i.- *Low-tech industries*: Final products in these industries remain basically the same through time. Production processes are relatively simple and therefore innovation efforts are focused more in marketing and distribution rather than in production. These industries have the lowest R&D intensity.

ii.- *Stable-tech industries*: Competition in these industries is based more on the improvement of existing product and production processes rather than in the development of new products. R&D efforts are more essential and critical than in low-tech industries, these efforts concentrating on process improvements and cost reductions, rather than on the development of new products.

iii.- *High-tech industries*: New-product development is the characteristic element of inter-firm competition. These industries have the highest R&D intensity. Innovation efforts must be devoted not only to improvement of existing products, but also to fund the high cost of developing new products. These industries also have the longest time horizons.

These technological classifications are also relevant to exit decisions, and are closely related to the market structure models mentioned above, where innovation plays a crucial role for firm survival.

3.5.5 Variables included

The source for firm-level information was *Standard & Poor's Compustat Industrial Annual.* Industry-level figures were obtained from the 1997 Economic Census's *Concentrating Ratios in Manufacturing* and *Manufacturing General Summary* reports. Unless otherwise noted, the term "Industry" refers to 4 and in some cases 6 digit level NAICS, each case depending on Compustat's available classification for each firm. The term "firm" refers to companies -as opposed to establishments-.

The variables used are the following:

i.- Scale variables:

$\frac{\texttt{Empl}}{\texttt{Avg.Empl}}$ =Ratio of firm's number of employees to the average number of employees in the firm's industry.

`Herfindahl`= Industry's Herfindahl index

`20-Largest Share`= Percent of value of shipments accounted for by the 20 largest companies in the industry

ii.- Market value variables

$\texttt{q}$= Tobin's 'q'.

$\texttt{eps}$= Earnings per-share.

$\frac{\texttt{Longterm debt}}{\texttt{Total assets}}$ = Ratio of firm's long term debt to its total assets.

iii.- Innovation-effort variables

$\frac{\texttt{Intangible assets}}{\texttt{Total assets}}$ = Ratio of firm's intangible assets to its total assets.

$\frac{\texttt{R\&D Expenditure}}{\texttt{Total sales}}$ = Ratio of firm's total R&D expenditure to its total sales.

iv.- Technological characteristics variables

Hall and Vopel (1997) have constructed a classification table for 4-digit SIC industries based on Chandler's technological segments. Using such a table, the following variable was included:

$$\texttt{Tech segment} = \begin{cases} 1 & \text{If firm belongs to low-tech segment} \\ 2 & \text{If firm belongs to stable tech-short horizon segment.} \\ 3 & \text{If firm belongs to stable tech-long horizon segment.} \\ 4 & \text{If firm belongs to hi-tech segment.} \end{cases}$$

3.6 Distributional assumptions

Let $\boldsymbol{Z} \equiv (\boldsymbol{X}_{R\&D}, \boldsymbol{X}_{Stay})$. Then, we will assume the following:

i.- $\boldsymbol{Z}$ have unknown joint cdf given by $G_{\boldsymbol{Z}}(\boldsymbol{z})$. We will denote the corresponding pdf by $dG_{\boldsymbol{Z}}(\boldsymbol{z})$.

ii.- Conditional on $\boldsymbol{Z}$, the shocks $\varepsilon_{R\&D}$ and ε_{Stay} have marginal cdfs given by $F_R(\epsilon_{R\&D})$ and $F_S(\epsilon_{Stay})$ respectively. Their joint cdf is given by $F_{\text{R, S}}(\epsilon_{R\&D}, \epsilon_{Stay})$. We will assume a particular functional form for these distributions, with parameters independent of $\boldsymbol{Z}$[2]

[2]This assumption can be relaxed by using the technique to deal with endogeneity detailed in section 4.4.8 above.

3.7 Information assumptions

We will assume the following:

1.- When firms make their optimal choices, the variables $\boldsymbol{X}_{R\&D}$, $\varepsilon_{R\&D}$, $\boldsymbol{X}_{Exit}$ and ε_{Exit} are privately known.[3]

2.- The variables $\boldsymbol{X}_{R\&D}$ and $\boldsymbol{X}_{Exit}$ become available (for the econometrician) some time after the optimal choices have been made. The variables $\varepsilon_{R\&D}$ and ε_{Exit} remain unknown to the econometrician.

To preserve the simultaneous nature of the game, we must assume that the R&D and exit decisions made in period t are irreversible for period $t+1$. It is clear that leaving or staying in business fits this commitment description.

The argument to model the decision to increase or decrease R&D expenditure as a pre-commitment could be based on the sunk-cost nature of R&D investments and most importantly, by noting that the returns to R&D investment are hard to measure in the short run. Because of this, it is unlikely that gradual availability of small pieces of information about other firms would result in a dramatic change in a particular firm's short-run R&D decision. Finally, one can also argue that in general, a firm's budget plan (made in advance) is rather rigid in the short run.

It is hard however, to justify the actual figure of R&D expenditure as a pre-commitment, that is why the model focuses on a qualitative R&D choice: increase or decrease R&D expenditure.

3.8 Beliefs and equilibrium conditions

As we mentioned above, when making their optimal choices, firms can't observe the population proportions $\pi^S_{R\&D}$, π^S_{Exit}, $\pi^L_{R\&D}$ and π^L_{Exit}. Firms will then maximize the expectation of their payoff functions (2) and (3). Let

$$E_i[\pi^S_{R\&D}] = \text{Firm i's expectation of } \pi^S_{R\&D}$$
$$E_i[\pi^S_{Exit}] = \text{Firm i's expectation of } \pi^S_{Exit}$$

[3] All that follows can be readily adapted to a case in which components of $\boldsymbol{X}_{R\&D}$ and $\boldsymbol{X}_{Exit}$ are public information, as long as the result is the creation of a finite number of types, depending on the realization of these publicly known variables.

and so on for $\pi^L_{R\&D}$ and π^L_{Exit}. In equilibrium, due to the informational assumptions of the model, all firms must have the same beliefs. Denote these beliefs as $\overline{\pi}^S_{R\&D}$, $\overline{\pi}^S_{Exit}$, $\overline{\pi}^L_{R\&D}$ and $\overline{\pi}^L_{Exit}$. Linearity of the payoff functions (2) and (3) allows us to simply plug in these beliefs instead of the true proportions to obtain the expected payoffs, which are given by:

For a small firm:

$$\begin{aligned} \overline{u}^S_{R\&D} &= \alpha^S_1 \overline{\pi}^S_{R\&D} + \alpha^S_2 \overline{\pi}^L_{R\&D} + \boldsymbol{\beta}'_{R\&D} \boldsymbol{X}_{R\&D} \\ \overline{u}^S_{Exit} &= \gamma^S_1 \overline{\pi}^S_{Exit} + \gamma^S_2 \overline{\pi}^L_{Exit} + \boldsymbol{\beta}'_{Stay} \boldsymbol{X}_{Stay} \end{aligned} \tag{4}$$

for a large firm:

$$\begin{aligned} \overline{u}^L_{R\&D} &= \alpha^L_1 \overline{\pi}^L_{R\&D} + \alpha^L_2 \overline{\pi}^S_{R\&D} + \boldsymbol{\beta}'_{R\&D} \boldsymbol{X}_{R\&D} \\ \overline{u}^L_{Exit} &= \gamma^L_1 \overline{\pi}^L_{Exit} + \gamma^L_2 \overline{\pi}^S_{Exit} + \boldsymbol{\beta}'_{Stay} \boldsymbol{X}_{Stay} \end{aligned} \tag{5}$$

for a large one.

A small firm will stay iff $\overline{u}^S_{Stay} > 0$ and will increase its R&D expenditure iff $\overline{u}^S_{R\&D}$. The same stay and R&D rules follow for a large firm, with $\overline{u}^L_{Stay}$ and $\overline{u}^L_{R\&D}$ respectively.

3.9 Estimation and Results

3.9.1 Identification

Identification concerns are very important in interactions-based models. This section examines the issues related to the proposed model. The issues addressed here are relevant in all potential applications. The specific implications and assumptions would depend on the particular model.

Payoff functions need normalization

It turns out that the parameters (α^S_1, α^S_2), (γ^S_1, γ^S_2), (α^L_1, α^L_2) and (γ^L_1, γ^L_2) can't be identified from the expected payoff functions as they are stated in (4) and (5). Some normalization is needed. I will assume that Small and Large firms' expected payoff functions can be expressed in the following symmetric way:

$$\begin{aligned} \overline{u}^S_{R\&D} &= \widetilde{\alpha}_1 \overline{\pi}^S_{R\&D} + \widetilde{\alpha}_2 \overline{\pi}_{R\&D} + \boldsymbol{\beta}'_{R\&D} \boldsymbol{X}_{R\&D} - \varepsilon_{R\&D} \\ \overline{u}^S_{Stay} &= \widetilde{\gamma}_1 \overline{\pi}^S_{Exit} + \widetilde{\gamma}_2 \overline{\pi}_{Exit} + \boldsymbol{\beta}'_{Stay} \boldsymbol{X}_{Stay} - \varepsilon_{Exit} \end{aligned} \tag{6}$$

$$\begin{aligned} \overline{u}^L_{R\&D} &= \widetilde{\alpha}_1 \overline{\pi}^L_{R\&D} + \widetilde{\alpha}_2 \overline{\pi}_{R\&D} + \boldsymbol{\beta}'_{R\&D} \boldsymbol{X}_{R\&D} - \varepsilon_{R\&D} \\ \overline{u}^L_{Stay} &= \widetilde{\gamma}_1 \overline{\pi}^L_{Exit} + \widetilde{\gamma}_2 \overline{\pi}_{Exit} + \boldsymbol{\beta}'_{Stay} \boldsymbol{X}_{Stay} - \varepsilon_{Exit} \end{aligned} \tag{7}$$

Where $\overline{\pi}_{R\&D}$ and $\overline{\pi}_{Exit}$ are the expectations of the overall population probabilities of increasing R&D expenditure (given that the firm stays) and leaving the industry respectively. That is:

$$\overline{\pi}_{R\&D} = \eta^L \overline{\pi}^L_{R\&D} + (1-\eta^L)\overline{\pi}^S_{R\&D} \quad \overline{\pi}_{Exit} = \eta^L \overline{\pi}^L_{Exit} + (1-\eta^L)\overline{\pi}^S_{Exit}$$

where η^L is the population probability of observing a Large firm. The original interaction parameters can be recovered as:

For a small firm:

$$\alpha^S_1 = \widetilde{\alpha}_1 + \widetilde{\alpha}_2(1-\eta^L), \quad \alpha^S_2 = \widetilde{\alpha}_2 \eta^L, \quad \gamma^S_1 = \widetilde{\gamma}_1 + \widetilde{\gamma}_2(1-\eta^L), \quad \gamma^S_2 = \widetilde{\gamma}_2 \eta^L$$

For a large firm:

$$\alpha^L_1 = \widetilde{\alpha}_1 + \widetilde{\alpha}_2 \eta^L, \quad \alpha^L_2 = \widetilde{\alpha}_2(1-\eta^L), \quad \gamma^L_1 = \widetilde{\gamma}_1 + \widetilde{\gamma}_2 \eta^L, \quad \gamma^L_2 = \widetilde{\gamma}_2(1-\eta^L)$$

The normalization chosen for (6) and (7) has the feature that the relative importance placed on the decisions of Small and Large firms depends -partially- on their relative presence in the population as reflected by η_L.

We will denote:

$$\begin{aligned} \boldsymbol{\theta}_1 &\equiv \left(\overline{\pi}^S_{R\&D}, \overline{\pi}^S_{Exit}, \overline{\pi}^L_{R\&D}, \overline{\pi}^L_{Exit}, \overline{\pi}_{R\&D}, \overline{\pi}_{Exit}, \eta^L\right) \\ \boldsymbol{\theta}_2 &\equiv \left(\widetilde{\alpha}_1, \widetilde{\alpha}_2, \widetilde{\gamma}_1, \widetilde{\gamma}_2, \boldsymbol{\beta}'_{R\&D}, \boldsymbol{\beta}'_{Stay}\right)' \\ \boldsymbol{\theta} &\equiv \left(\boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2\right)' \end{aligned}$$

Now let:

$$\begin{aligned} \delta^S_{R\&D}(\boldsymbol{\theta}, \boldsymbol{Z}) &\equiv \widetilde{\alpha}_1 \overline{\pi}^S_{R\&D} + \widetilde{\alpha}_2 \overline{\pi}_{R\&D} + \boldsymbol{\beta}'_{R\&D} \boldsymbol{X}_{R\&D} \\ \delta^S_{Stay}(\boldsymbol{\theta}, \boldsymbol{Z}) &\equiv \widetilde{\gamma}_1 \overline{\pi}^S_{Exit} + \widetilde{\gamma}_2 \overline{\pi}_{Exit} + \boldsymbol{\beta}'_{Stay} \boldsymbol{X}_{Stay} \\ \delta^L_{R\&D}(\boldsymbol{\theta}, \boldsymbol{Z}) &= \widetilde{\alpha}_1 \overline{\pi}^L_{R\&D} + \widetilde{\alpha}_2 \overline{\pi}_{R\&D} + \boldsymbol{\beta}'_{R\&D} \boldsymbol{X}_{R\&D} \\ \delta^L_{Stay}(\boldsymbol{\theta}, \boldsymbol{Z}) &= \widetilde{\gamma}_1 \overline{\pi}^L_{Exit} + \widetilde{\gamma}_2 \overline{\pi}_{Exit} + \boldsymbol{\beta}'_{Stay} \boldsymbol{X}_{Stay} \end{aligned}$$

Conditional on $\boldsymbol{Z}$ we have the following:

For a small firm:

$Pr\big(\text{Increasing R\&D given that firm stays}\big) \equiv Pr^S(R\&D \mid Stay, \boldsymbol{Z}, \boldsymbol{\theta}) = \frac{F_{\text{R, S}}(\delta^S_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}),\delta^S_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))}{F_{\text{S}}(\delta^S_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))}$

$Pr\big(\text{Leaving the industry}\big) \equiv Pr^S(Exit \mid \boldsymbol{Z}, \boldsymbol{\theta}) = 1 - F_{\text{S}}(\delta^S_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))$

For a large firm:

$Pr\big(\text{Increasing R\&D given that firm stays}\big) \equiv Pr^L(R\&D \mid Stay, \boldsymbol{Z}, \boldsymbol{\theta}) = \frac{F_{\text{R, S}}(\delta^L_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}),\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))}{F_{\text{L}}(\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))}$

$Pr\big(\text{Leaving the industry}\big) \equiv Pr^L(Exit \mid \boldsymbol{Z}, \boldsymbol{\theta}) = 1 - F_{\text{L}}(\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))$

Because a firm's size is a determinant of R&D and exit choices, it will be included in $\boldsymbol{Z}$. For simplicity, denote size as $Z_1 \in \boldsymbol{Z}$. We will assume that a firm is "Large" if $Z_1 > k_L$ for some cut-off value k_L and "Small" otherwise.

Define:

$$
\begin{aligned}
\psi_1(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \overline{\pi}^S_{R\&D} - \frac{\mathbf{1}\{Z_1 \le k_L\} Pr^S(R\&D \mid Stay, \boldsymbol{Z}, \boldsymbol{\theta})}{1-\eta^L}\\
\psi_2(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \overline{\pi}^L_{R\&D} - \frac{\mathbf{1}\{Z_1 > k_L\} Pr^L(R\&D \mid Stay, \boldsymbol{Z}, \boldsymbol{\theta})}{\eta^L}\\
\psi_3(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \overline{\pi}^S_{Exit} - \frac{\mathbf{1}\{Z_1 \le k_L\} Pr^S(Exit \mid \boldsymbol{Z}, \boldsymbol{\theta})}{1-\eta^L}\\
\psi_4(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \overline{\pi}^L_{Exit} - \frac{\mathbf{1}\{Z_1 > k_L\} Pr^L(Exit \mid \boldsymbol{Z}, \boldsymbol{\theta})}{\eta^L}\\
\psi_5(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \overline{\pi}_{R\&D} - \Big[\mathbf{1}\{Z_1 \le k_L\} Pr^S(R\&D \mid Stay, \boldsymbol{\theta}, \boldsymbol{Z}) + \mathbf{1}\{Z_1 > k_L\} Pr^L(R\&D \mid Stay, \boldsymbol{\theta}, \boldsymbol{Z})\Big]\\
\psi_6(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \overline{\pi}_{Exit} - \Big[\mathbf{1}\{Z_1 \le k_L\} Pr^S(Exit \mid \boldsymbol{Z}, \boldsymbol{\theta}) + \mathbf{1}\{Z_1 > k_L\} Pr^L(Exit \mid \boldsymbol{Z}, \boldsymbol{\theta})\Big]\\
\psi_7(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \eta^L - \mathbf{1}\{Z_1 > k_L\}\\
\boldsymbol{\Psi}(\boldsymbol{\theta},\boldsymbol{Z}) &\equiv \Big(\psi_1(\boldsymbol{\theta},\boldsymbol{Z}), \psi_2(\boldsymbol{\theta},\boldsymbol{Z}), \psi_3(\boldsymbol{\theta},\boldsymbol{Z}), \psi_4(\boldsymbol{\theta},\boldsymbol{Z}), \psi_5(\boldsymbol{\theta},\boldsymbol{Z}), \psi_6(\boldsymbol{\theta},\boldsymbol{Z}), \psi_7(\boldsymbol{\theta},\boldsymbol{Z})\Big)'
\end{aligned}
$$

Then, *in a Bayesian Nash equilibrium beliefs must satisfy*:

$$\int_{\boldsymbol{z}} \boldsymbol{\Psi}(\boldsymbol{\theta},\boldsymbol{z}) dG_{\boldsymbol{Z}}(\boldsymbol{z}) = \mathbf{0} \tag{8}$$

Existence of equilibria

For a given value of $\boldsymbol{\theta}_2$, we're interested in knowing if there exists a set of beliefs $\boldsymbol{\theta}_1$ such that the equilibrium condition (8) is satisfied. A sufficient condition for the existence of equilibria is that the joint and marginal distributions of $\varepsilon_{R\&D}$ and ε_{Exit} be continuous. Existence of equilibria for an arbitrary value of $\boldsymbol{\theta}_2$ follows from Brower's fixed point theorem. Therefore, an equilibrium must exist for $\boldsymbol{\theta}_2^0$, the true population values of $\boldsymbol{\theta}_2$. Details are given in the appendix.

Uniqueness of equilibrium

The question of *uniqueness* of equilibrium is a very important one. If, for the true values of $\boldsymbol{\theta}_2$ there exists more than one set of beliefs $\boldsymbol{\theta}_1$ that satisfy equilibrium condition (8), then we would have to make additional assumptions about which, among the set of equilibrium beliefs is used by each firm. In our formulation for example, we would have to assume that all firms use the same equilibrium beliefs. Alternative formulations could have all small firms having the same equilibrium beliefs and all large ones also having the same equilibrium beliefs, which could be different from those used by the small firms.

The question of uniqueness can be analyzed by looking at the jacobian:

$$\nabla_{\boldsymbol{\theta}_1} \int_{\boldsymbol{z}} \Psi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0, \boldsymbol{z}) dG_{\boldsymbol{Z}}(\boldsymbol{z})$$

where as before, $\boldsymbol{\theta}_2^0$ represent the true population values of $\boldsymbol{\theta}_2$.

Local uniqueness will be guaranteed if the jacobian:

$$\nabla_{\boldsymbol{\theta}_1} \int_{\boldsymbol{z}} \Psi(\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^0, \boldsymbol{z}) dG_{\boldsymbol{Z}}(\boldsymbol{z}), \text{ where } \boldsymbol{\theta}_1^* \text{ is a solution to } \int_{\boldsymbol{z}} \boldsymbol{\Psi}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0, \boldsymbol{z}) dG_{\boldsymbol{Z}}(\boldsymbol{z}) = \mathbf{0}$$

has rank equal to seven (full rank). An example of a *sufficient* condition for global uniqueness would be to assume that the jacobian $\nabla_{\boldsymbol{\theta}_1} \int_{\boldsymbol{z}} \Psi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0, \boldsymbol{z}) dG_{\boldsymbol{Z}}(\boldsymbol{z})$ has either: (i) only strictly positive principal minors or (ii) only strictly negative principal minors for all $\boldsymbol{\theta}_1 \in [0,1]^7$. This is a version of the *Gale-Nikaido theorem* that guarantees that $\int_{\boldsymbol{z}} \boldsymbol{\Psi}(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0 \boldsymbol{z}) dG_{\boldsymbol{Z}}$ is a one-to-one function of $\boldsymbol{\theta}_1$ and therefore, that the equilibrium is unique. Simply put, it says that the jacobian not only has to be non-singular, but it also has to remain either positive quasi definite or negative quasi definite for all values of $\boldsymbol{\theta}_1$.

Distinction between small and large firms has to be statistically relevant, and equilibrium beliefs must be non-degenerate:

Existence and uniqueness of equilibrium have to do with the identification of $\boldsymbol{\theta}_1$, the vector of beliefs. The functional form assumed for the expected payoff functions requires two additional conditions for the identification of $\boldsymbol{\theta}_2$. Basically, these conditions are necessary for the asymptotic invertibility of the Hessian for the first order conditions satisfied by the EL estimator[4]:

(i) All equilibrium beliefs $\boldsymbol{\theta}_1^0$ must be strictly between zero and one. That is, in equilibrium the population probability of choosing any of the three possible actions must be strictly positive and this must hold for both small and large firms. This is a *necessary* condition for identification of $\boldsymbol{\theta}_2$.

(ii) The conditional distributions $G(\boldsymbol{Z} \mid Z_1 > k_L)$ and $G(\boldsymbol{Z} \mid Z_1 \leq k_L)$ are *not identical.* This means that there is a (statistically) meaningful difference between small and large firms. This is a sufficient condition for general values of $\boldsymbol{\theta}_1^0$, but it becomes a necessary one for some nontrivial possible values of $\boldsymbol{\theta}_1^0$.

Conditions (i) and (ii) together simply require that the proposed interaction be meaningful: if (i) is violated, then it would be common knowledge for example, that all small firms choose the same action (they all leave the industry, for example). If (ii) is violated, it would imply that there is no strategic interaction that takes place in the "size" dimension: there is nothing essentially different between small and large firms. Violations to (i) or (ii) seem implausible in reality.

3.9.2 Estimation

Conditional likelihood

Having dealt with identification issues, we now present the estimator. First, let

$$\boldsymbol{Y} = (Y_1, Y_2, Y_3)$$

[4]The role played by these identification conditions is parallel to the one played by the conditions necessary to assume invertibility of the information matrix in the usual maximum likelihood estimation problems.

where Y_1, Y_2 and Y_3 were defined in section 3.1.1. Then:

For a *small* firm, the conditional log-likelihood of $\boldsymbol{Y}$ given $\boldsymbol{Z}$ is

$$\log f^S(\boldsymbol{Y} \mid \boldsymbol{Z},\boldsymbol{\theta}) = Y_1 \log F_{\text{R, S}}(\delta^S_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}), \delta^S_{Stay}(\boldsymbol{\theta},\boldsymbol{Z})) + Y_3 \log\left(1 - F_{\text{S}}(\delta^S_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))\right) + Y_2 \log\left(F_{\text{S}}(\delta^S_{Stay}(\boldsymbol{\theta},\boldsymbol{Z})) - F_{\text{R, S}}(\delta^S_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}), \delta^S_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))\right)$$

and for a *large* firm, it's given by:

$$\log f^L(\boldsymbol{Y} \mid \boldsymbol{Z},\boldsymbol{\theta}) = Y_1 \log F_{\text{R, S}}(\delta^L_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}), \delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z})) + Y_3 \log\left(1 - F_{\text{S}}(\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))\right) + Y_2 \log\left(F_{\text{S}}(\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z})) - F_{\text{R, S}}(\delta^L_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}), \delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))\right)$$

Therefore, for an arbitrary firm, the log-likelihood is naturally

$$\log f(\boldsymbol{Y} \mid \boldsymbol{Z},\boldsymbol{\theta}) = \mathbf{1}\{Z_1 \leq k_L\} \log f^L(\boldsymbol{Y} \mid \boldsymbol{Z},\boldsymbol{\theta}) + \mathbf{1}\{Z_1 > k_L\} \log f^S(\boldsymbol{Y} \mid \boldsymbol{Z},\boldsymbol{\theta})$$

Empirical Likelihood estimator

The proposed empirical Likelihood estimator $\widehat{\boldsymbol{\theta}}^{\text{EL}}$ is the solution to:

$$\begin{aligned} &\underset{\boldsymbol{\theta},\{p_i\}_{i=1}^N}{\text{Maximize}} \sum_{i=1}^N \log f(\boldsymbol{y}_i \mid \boldsymbol{z}_i,\boldsymbol{\theta}) + \sum_{i=1}^N \log p_i \\ &\text{subject to} \quad p_i \geq 0, \quad \sum_{i=1}^N p_i = 1, \quad \sum_{i=1}^N \Psi(\boldsymbol{z}_i,\boldsymbol{\theta}) = 0 \end{aligned} \tag{9}$$

It should be pointed out that this methodology is closely related to the one proposed by Cosslett (1981), who proposed a pseudo maximum likelihood estimator for discrete choice models with choice based sampling in which he estimated the model's parameters along with a set of weights subject to the same "well-behaved pdf" conditions presented here and subject to the probabilities of sampling from the different subpopulations or strata. He showed that the resulting estimator was efficient. Cosslett's proposal wasn't formally embedded in the empirical likelihood methodology which wasn't proposed and studied until 7 years later.

The asymptotic properties of $\widehat{\boldsymbol{\theta}}^{\text{EL}}$ are detailed in the appendix. Some of its most important properties are:

1.- $\widehat{\boldsymbol{\theta}}^{\text{EL}}$ has the same asymptotic distribution as the efficient GMM estimator based on the moment conditions

$$E\big[\nabla_{\boldsymbol{\theta}} \log f(\boldsymbol{Y} \mid \boldsymbol{Z},\boldsymbol{\theta}_0)\big] = 0 \quad \text{and} \quad E\big[\Psi(\boldsymbol{Z},\boldsymbol{\theta}_0)\big] = 0$$

This result is parallel to that of Imbens (1992), where he proposed a GMM estimator that had the same efficiency as Cosslett's (1981) estimator for choice-based sampling. As it was mentioned previously, EL and GMM estimators typically have the same asymptotic distribution to first order of approximation. However, as it was also mentioned above, the superior properties of EL for small samples has been well documented. EL has a smaller variance to second order of approximation. Computationally it's also more convenient since no first-step estimators and no weight matrix have to be computed.

2.- $\widehat{\boldsymbol{\theta}}^{\text{EL}}$ is more efficient than the estimator that solves

$$\underset{\boldsymbol{\theta}}{\text{Maximize}} \sum_{i=1}^{N} \log f(\boldsymbol{y}_i \mid \boldsymbol{z}_i,\boldsymbol{\theta}) \quad \text{subject to} \quad \frac{1}{N}\sum_{i=1}^{N} \Psi(\boldsymbol{z}_i,\boldsymbol{\theta}) = 0$$

i.e, the one that also uses the equilibrium conditions but imposes the uniform weights $1/N$. This shows the importance of simultaneously estimating the unknown pdf $dG(\boldsymbol{Z})$ and the parameters of interest.

3.- Using additional available information about the population distribution of $\boldsymbol{Z}$ increases the efficiency of $\widehat{\boldsymbol{\theta}}^{\text{EL}}$. We'd have this additional information for example, if we know that $E[Z_\ell] = c_0$ for some $Z_\ell \in \boldsymbol{Z}$ and we know the value of c_0, or if we know that $Pr(Z_\ell \leq q_0) = 0.5$, for a known q_0. More importantly, we can also show that these efficiency gains are also approximately true if we use census information or another target population of size M, independent from our sample as long as N/M is relatively small. Hellerstein and Imbens (1999) proved a very similar result for a different kind of EL estimation. Imbens and Lancaster (1994) showed how census statistics could be approximately considered as population moments for GMM problems.

The model was estimated using different combinations of the explanatory variables listed in section 3.5.5. Because size criteria were derived from 1997 census figures, only cross sections for years close to 1997 were included. The model was estimated separately for the cross sections of 1997, 1998 and 1999. The total number of firms in 1997 was 2139. Approximately 7.4, 9.4 and 8.6 percent of firms left in 1997, 1998 and 1999 respectively. Approximately 59.2, 48.5 and 56.3 percent of firms chose to stay and increase R&D expenditure for the following year in 1997, 1998 and 1999 respectively. Finally, "large" firms made up approximately 54.3, 58.0 and 56.3 percent of the sample in each of the three years.

The chosen distributions for $\varepsilon_{R\&D}$ and ε_{Stay} were:

$$F_S(\epsilon_{Stay}) = \frac{1}{1+e^{-\epsilon_{\text{Stay}}}}, \quad F_R(\epsilon_{R\&D}) = \frac{1}{1+e^{-\epsilon_{\text{R\&D}}}}, \quad \text{(i.e, both logistic)},$$

$$F_{S,R}(\epsilon_{Stay}, \epsilon_{R\&D}) = \frac{F_S(\epsilon_{Stay})F_R(\epsilon_{R\&D})}{1-\sigma(1-F_S(\epsilon_{Stay}))(1-F_R(\epsilon_{R\&D}))}, \quad -1 \leq \sigma \leq 1 \tag{10}$$

The joint distribution uses the Ali-Mikhail-Haq copula. This copula includes Gumbel's bivariate logistic distribution as a special case when $\sigma = 1$.[5]

Unconstrained MLE: The empirical likelihood ratio test plays a very important role in the estimation. For a given value of σ, let $\ell^{\text{MLE}}(\widehat{\boldsymbol{\theta}}^{\text{MLE}} \mid \sigma)$ be the maximum likelihood of the unconstrained model, which is estimated by replacing (6) and (7) with the equations:

$$\begin{aligned} \overline{u}^S_{R\&D} &= \zeta^S_{R\&D} + \boldsymbol{\beta}'_{R\&D}\boldsymbol{X}_{R\&D} - \varepsilon_{R\&D} \\ \overline{u}^S_{Stay} &= \zeta^S_{Stay} + \boldsymbol{\beta}'_{Stay}\boldsymbol{X}_{Stay} - \varepsilon_{Exit} \end{aligned} \tag{6'}$$

and

$$\begin{aligned} \overline{u}^L_{R\&D} &= \zeta^L_{R\&D} + \boldsymbol{\beta}'_{R\&D}\boldsymbol{X}_{R\&D} - \varepsilon_{R\&D} \\ \overline{u}^L_{Stay} &= \zeta^L_{Stay} + \boldsymbol{\beta}'_{Stay}\boldsymbol{X}_{Stay} - \varepsilon_{Exit} \end{aligned} \tag{7'}$$

No Bayesian-Nash equilibrium constraints are imposed and the distributions are those described in (10). For a given σ, unconstrained MLE only needs to estimate $\zeta^S_{R\&D}$, ζ^S_{Stay}, $\zeta^L_{R\&D}$, ζ^L_{Stay}, $\boldsymbol{\beta}_{R\&D}$ and $\boldsymbol{\beta}_{Stay}$.

Estimation algorithm

[5] Copulas are functions that join multivariate distribution functions to their one-dimensional margins. See Nelsen (1999).

A grid search in $[-1,1]$ was done for σ. For each value of σ, let $\ell^{\mathrm{EL}}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}} \mid \sigma)$ be the parametric portion of the maximum empirical likelihood. Let $\{p_i(\sigma)\}_{i=1}^N$ be the optimal weights. For each σ, the parameter vector $\boldsymbol{\theta}$ was estimated.

For each σ, the empirical likelihood ratio

$$\mathcal{R}(\sigma) = 2 \times \left\{ \ell^{\mathrm{MLE}}(\widehat{\boldsymbol{\theta}}^{\mathrm{MLE}} \mid \sigma) + \sum_{i=1}^{N} \log(1/N) - \ell^{\mathrm{EL}}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}} \mid \sigma) - \sum_{i=1}^{N} p_i(\sigma) \right\}$$

was computed. The goal of the grid search for σ was to approximate the one that minimized $\mathcal{R}(\sigma)$. This was done for a number of combinations of explanatory variables $\boldsymbol{X}_{R\&D}$ and $\boldsymbol{X}_{Stay}$, and different values of k_L -the cutoff value to determine if a firm is "Large" or "Small"-. The following set of variables consistently yielded the best results for 1997, 1998 and 1999:

$$\boldsymbol{X}_{R\&D} = \{ \tfrac{\texttt{Empl}}{\texttt{Avg.Empl}}, \texttt{eps}, \tfrac{\texttt{Intangible assets}}{\texttt{Total assets}}, \texttt{q}, \texttt{Techsegment} \}$$

$$\boldsymbol{X}_{Stay} = \{ \tfrac{\texttt{Empl}}{\texttt{Avg.Empl}}, \texttt{eps}, \tfrac{\texttt{Longterm debt}}{\texttt{Total assets}}, \texttt{q}, \texttt{Techsegment} \}$$

The variable "Size", denoted by Z_1 in conditions (8) was $\frac{\texttt{Empl}}{\texttt{Avg.Empl}}$. For each σ, the likelihood ratio $\mathcal{R}(\sigma)$ was rather stable for cutoff values k_L close to 2. Therefore, a firm was considered "large" if $\frac{\texttt{Empl}}{\texttt{Avg.Empl}} > 2$. The statistic $\mathcal{R}(\sigma)$ was very small for the three years. However, it was remarkably smaller for 1998. All Lagrange multipliers were statistically equal to zero for the three years, but the likelihood ratio statistic $\mathcal{R}(\sigma)$ was far smaller for 1998. The following section discusses the results.

3.9.3 A digression on endogeneity

A very important stochastic assumption was the independence between $\varepsilon_{R\&D}$ and $\boldsymbol{X}_{R\&D}$ and between ε_{Exit} and $\boldsymbol{X}_{Exit}$. Independence assumptions are difficult to defend and often objectionable in econometric models. It turns out that all the asymptotic results for $\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}$ still hold if *all* we assume is that:

For a small firm:

$$\begin{aligned}
E[Y_1 \mid \boldsymbol{Z}, \boldsymbol{\theta}] &= F_{\mathrm{R,\,S}}(\delta^S_{R\&D}(\boldsymbol{\theta}, \boldsymbol{Z}), \delta^S_{Stay}(\boldsymbol{\theta}, \boldsymbol{Z})) \\
E[Y_2 \mid \boldsymbol{Z}, \boldsymbol{\theta}] &= F_{\mathrm{S}}(\delta^S_{Stay}(\boldsymbol{\theta}, \boldsymbol{Z})) - F_{\mathrm{R,\,S}}(\delta^S_{R\&D}(\boldsymbol{\theta}, \boldsymbol{Z}), \delta^S_{Stay}(\boldsymbol{\theta}, \boldsymbol{Z})) \\
E[Y_3 \mid \boldsymbol{Z}, \boldsymbol{\theta}] &= 1 - F_{\mathrm{S}}(\delta^S_{Stay}(\boldsymbol{\theta}, \boldsymbol{Z}))
\end{aligned}$$

and for a large firm:

$$
\begin{aligned}
E[Y_1 \mid \boldsymbol{Z},\boldsymbol{\theta}] &= F_{\mathrm{R,\,S}}(\delta^L_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}),\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))\\
E[Y_2 \mid \boldsymbol{Z},\boldsymbol{\theta}] &= F_{\mathrm{S}}(\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z})) - F_{\mathrm{R,\,S}}(\delta^L_{R\&D}(\boldsymbol{\theta},\boldsymbol{Z}),\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))\\
E[Y_3 \mid \boldsymbol{Z},\boldsymbol{\theta}] &= 1 - F_{\mathrm{S}}(\delta^L_{Stay}(\boldsymbol{\theta},\boldsymbol{Z}))
\end{aligned}
$$

and if we assume that these conditional expectations are common knowledge among all firms. These modified stochastic and informational assumptions generate the exact same equilibrium conditions, estimators and asymptotic results. We have to keep in mind that the primary goal of our models is to estimate the interaction parameters, estimation of the parameters of $\boldsymbol{Z}$ is secondary.

The cost of using these assumptions is of course, that we have to drop the payoff maximization story. The way beliefs enter the functional forms for these conditional expectations could also seem "arbitrary", while it was completely transparent in the original formulation of the model. However, the functional form for the original payoff functions was also arbitrary. More importantly, the Bayesian-Nash equilibrium conditions do not change. It seems wiser to use these as the basic assumptions of the model. Endogeneity concerns have always been more daunting to econometricians than functional misspecification of conditional moments.

If for some reason endogeneity concerns are overwhelming and we wanted to hold on to the payoff maximization story, then we can still apply the methodology proposed here by adding a vector of exogenous variables $\widetilde{\boldsymbol{Z}} \in \mathbb{R}^s$ and assuming for example, that

$$\boldsymbol{Z} = \boldsymbol{\Pi}\widetilde{\boldsymbol{Z}} + \boldsymbol{v}$$

where $\boldsymbol{\Pi}$ is a $k \times s$ matrix. We will assume that $\widetilde{\boldsymbol{Z}}$ is independent of $\boldsymbol{v}$, $\varepsilon_{R\&D}$ and ε_{Exit}. We will denote its unknown pdf by $dG_{\widetilde{\boldsymbol{Z}}}(\widetilde{\boldsymbol{z}})$.

Endogeneity of $\boldsymbol{Z}$ could be assumed to take place through a dependence between $\boldsymbol{v}$ and $\varepsilon_{R\&D}$ and ε_{Exit}. We could then take the following steps:

1.- Assume a functional form for the distribution of $\boldsymbol{v}$. Denote it as $F_{\boldsymbol{v}}(\boldsymbol{v})$. Use it to obtain the conditional distribution of $\boldsymbol{Z}$ given $\widetilde{\boldsymbol{Z}}$. Denote the latter by $G_{\boldsymbol{Z}|\widetilde{\boldsymbol{Z}}}(\boldsymbol{Z} \mid \widetilde{\boldsymbol{Z}})$.

2.- Assume a functional form for the conditional distributions of $\varepsilon_{R\&D}$ given $\boldsymbol{v}$ and of ε_{Exit} given $\boldsymbol{v}$. Denote these by $F_{\varepsilon_{R\&D}|\boldsymbol{v}}(\epsilon_{R\&D} \mid \boldsymbol{v})$ and $F_{\varepsilon_{Exit}|\boldsymbol{v}}(\epsilon_{Exit} \mid \boldsymbol{v})$ respectively.

3.- Re-express the equilibrium condition and the log-likelihood in terms of $dF_{\boldsymbol{v}}(\boldsymbol{v})$, $G_{\boldsymbol{Z}|\widetilde{\boldsymbol{Z}}}(\boldsymbol{Z} \mid \widetilde{\boldsymbol{Z}})$, $F_{\varepsilon_{R\&D}|\boldsymbol{v}}(\epsilon_{R\&D} \mid \boldsymbol{v})$, $F_{\varepsilon_{Exit}|\boldsymbol{v}}(\epsilon_{Exit} \mid \boldsymbol{v})$ and the unknown $dG_{\widetilde{\boldsymbol{Z}}}(\widetilde{\boldsymbol{z}})$.

4.- Perform EL estimation with the re-expressed log-likelihood and equilibrium conditions. We would now have to estimate -in addition to $\boldsymbol{\theta}$- the parameters $\boldsymbol{\Pi}$ and all the possible unknown parameters of the distributions $dF_{\boldsymbol{v}}(\boldsymbol{v})$, $G_{\boldsymbol{Z}|\widetilde{\boldsymbol{Z}}}(\boldsymbol{Z} \mid \widetilde{\boldsymbol{Z}})$, $F_{\varepsilon_{R\&D}|\boldsymbol{v}}(\epsilon_{R\&D} \mid \boldsymbol{v})$ and $F_{\varepsilon_{Exit}|\boldsymbol{v}}(\epsilon_{Exit} \mid \boldsymbol{v})$. The optimal weights would now correspond to the unknown $\{dG_{\widetilde{\boldsymbol{Z}}}(\widetilde{\boldsymbol{z}}_i)\}_{i=1}^{N}$

3.9.4 Estimation Results

Results can be found in tables 1-3 in the appendix. We have the following main results:

1.- *Strategic interaction considerations are more important for R&D choices than for Exit choices:* The estimates for $(\alpha_1^S, \alpha_2^S, \alpha_2^L)$ were statistically significant at a 0.95 confidence level for all three years. The same was true for α_1^L for 1997 and 1999[6]. On the other hand, (γ_1^S, γ_1^L) were not significant for any of the three years, while (γ_2^S, γ_2^L) were significant only for 1998 and 1999, which shows that strategic considerations were not completely absent from firms' "exit" decision.

This is an important result since it is consistent with the assertion that firms are interested in anticipating R&D decisions made by others. Innovation has long-run implications for a firm's performance. On the other hand, the decision to leave the industry seems to be driven more by non-strategic considerations: it is hard for an expected change in market structure to determine a firm's decision to remain or to leave an industry. It seems that firms know that innovative firms tend to thrive regardless of the remaining number of firms in the industry: an industry's expected future size is a poor indicative by itself of future performance for a firm.

[6] α_1^L was statistically significant at approximately a 0.92 level for 1998.

2.- *R&D strategic behavior is competitive among firms of the same type and free rider-like among firms of different type:* For all three years we have that α_1^S and α_2^S are positive, and negative, respectively. The same is true for α_1^L and α_2^L. This is consistent with the assertion that firms behave competitively in terms of R&D with firms of their same type, but have a free-rider-like behavior with respect to firms of the opposite type. This effect is more significant for small firms than for large firms: $|\alpha_2^S| > |\alpha_2^L|$ consistently for the three years examined, which means that the free-rider behavior is more significant for small firms. Innovation by a given firm benefits all firms in the long run, generating a free-rider incentive, but it also gives that firm an advantage relative to its direct competitors. The results obtained are consistent with this assertion if we believe that firms of the same size tend to compete more fiercely with each other.

3.- *Firms don't seem to place any strategic importance on the exit decisions of firms of their same 'type', but they tend to care about the exit decisions of firms of a different 'type':* The estimates for γ_1^S and γ_1^L are never significant. This would be consistent with the assertion that competition among firms of the same type is a complex one: it involves far more than simply their decisions to stay or leave. We already concluded that this competition involves strategic R&D considerations. Once again: having fewer large firms in the industry doesn't assure a future positive performance for a given large firm: this performance is (at least partially) determined by the firm's ability to innovate. We also have that (γ_2^S, γ_2^L) are positive and significant for 1998 and 1999. Combined with result 2, this implies that firms "like it" when firms of a different type innovate, but they also "like it" when those firms leave. This is also consistent with the assertion that firms of different types compete in dimensions *other* than R&D (where they benefit from each other's R&D expenditure).

4.- *Expected actions taken by firms of the opposite 'type' tend to have a statistically greater impact than expected actions taken by firms of the same 'type':* There is no case -for any of the three years- in which α_1^S is statistically significant while α_2^S is not. The same is true for all other pairs of interaction parameters. We can see however, that γ_2^S and γ_2^L are significant for 1998 and 1999 while γ_1^S and γ_1^L are never significant. This is consistent with the assertion that firms tend to put more weight on the expected behavior of firms that are essentially different from them: The model is a static one, but firms may know more about the future behavior of those firms similar to them than about the future behavior of firms different from them. The model also examines only two decisions: to stay and to increase R&D expenditure; firms may give relatively less weight to firms similar to them because they know more about choices other than 'stay' and 'increase R&D'. On the other hand, firms may give relatively more weight to firms different from them because they know little about other choices made by them.

5.- *Strategic considerations seem to be more important for small firms than for large firms:* All strategic coefficients are statistically more significant for small firms than for large firms. Other things being equal, small firms' subsistence tends to be more precarious than that of large firms and they tend to be more interested in predicting future behavior of others. This seems consistent with the assertion that small firms tend to act more as "strategic followers" than large firms.

6.- *Estimated optimal weights show difference between small and large firms:* Identification required that the distinction between small and large firms be meaningful: the conditional distribution of $G(\boldsymbol{Z} \mid Z_1 > k_L)$ and $G(\boldsymbol{Z} \mid Z_1 \leq k_L)$ should not be identical. The estimated optimal weights tend to support this condition. This is shown in the appendix, where optimal weights are shown for 1997. When firms are ordered randomly, we see no pattern in the optimal weights, but when firms are ordered according to their size we can see that the optimal weights have a different pattern for small and large firms. The former seem to retain a uniform-like shape, while the latter become more volatile, non-uniform shape. This pattern was also present for 1998 and 1999.

Some comments on the estimation results for $\boldsymbol{\beta}_{R\&D}$ and $\boldsymbol{\beta}_{Exit}$:

1.- Earnings per share (EPS) were more significant than Tobin's 'q' among the financial variables included. Tobin's q seems to be more significant for the exit decision (1998 and 1999) than for the R&D decision. This was to be expected since Tobin's 'q' is a measure of the present value of a firm's investments. It is therefore a guidance of a firm's future viability. We have on the other hand, that firms' R&D efforts seem to respond significantly to changes in EPS. This is consistent with the innovation models mentioned previously, where the market values firms' innovation efforts: higher EPS tend to favor a firm's market value, and there is evidence that firms try to reinforce that effect with higher R&D expenditure.

2.- A firm's relative size in its industry is more relevant for exit decisions than for R&D decisions. The model consistently showed that both small and large firms care about R&D activities. A firm's size however, does have a larger influence over its market power and its ability to compete with other firms. Results show that relatively larger firms tend to have a higher probability of staying.

3.- A firm's technological category is relevant for R&D decisions. This is completely consistent with Chandler's description of technological segments for industries: firms in low-tech industries have less incentive than firms in hi-tech industries to increase their R&D expenditure. The ability to innovate is crucial for the survival of a hi-tech firm. There wasn't a significant relationship between a firm's decision to exit and its technological group.

4 Concluding Remarks

Asymmetric information is the appropriate setting for a number of interactions-based models. This asymmetric information exists because players can't observe (at least some of) the variables that determine other players' payoffs and therefore, their choices. Econometric estimation of these models entails the estimation of players' *beliefs* which are almost always unobservable. Using proxy variables for these beliefs is not a satisfactory answer to the problem. However, assuming that the observed behavior is the result of a Bayesian-Nash equilibrium implies that these beliefs must satisfy a set of clear-cut conditions. These conditions involve the unknown distribution of the privately observed variables. In a number of cases, portions of these privately observed variables may become available to the econometrician *after* the game was played.

In this case, estimation seems almost suited for empirical likelihood methods. This allows us to estimate *simultaneously* the payoff parameters, the beliefs and the unknown distribution of the privately-observed-available-afterwards-to-the-econometrician variables. Such an estimator was proposed, and its main properties were mentioned. Most importantly, the vast literature on

EL shows that it has better small sample properties than GMM -which could also be used for these models- it is also computationally more convenient: no first step estimators or weight matrix are needed. Identification issues are very important for general interactions-based models and they are also important for their asymmetric information counterparts. Issues such as existence and uniqueness of equilibrium are important, and thankfully more tractable than they are in general, perfect information models.

An application for an R&D model was analyzed and estimated. In a model where firms first decide to stay of exit their industry and then they decide to increase or decrease their R&D expenditure, we found evidence that strategic considerations are more important for R&D decisions than for exit decisions. Dividing firms in two types: "small" and "large" we found evidence that there is intense R&D strategic competition between firms of the same type, and at the same time free-rider-like strategic behavior between firms of different types. We also found that small firms are more eager than large ones to predict other firms' behavior. They tend to act more as "strategic followers".

The next step is to adapt this methodology to deal with dynamic models. The R&D model presented here seems naturally more for a dynamic formulation. To make this transition successfully, results from empirical likelihood for time series must be used. Empirical weights would now correspond to (overlapping) blocks of observations through time. There is no clear-cut criteria for the length of these blocks or their overlapping. Asymptotic results are a bit tricker. Empirical Likelihood for time series is an area in an early stage of development. However, it seems the natural way to deal with dynamic models of asymmetric information.

Appendix 1: Some proofs

Existence of equilibria:
To prove the existence of a solution to (8), note that the equilibrium conditions

$$\int_{\boldsymbol{z}} \Psi(\boldsymbol{\theta}, \boldsymbol{z}) dG(\boldsymbol{z}) = 0$$

can be expressed as

$$\overline{\pi}^S_{R\&D} = \int_{\boldsymbol{z}} \frac{\mathbf{1}\{z_1 \leq k_L\} Pr^S(R\&D \mid Stay, \boldsymbol{z}, \boldsymbol{\theta})}{1-\eta^L} dG_{\boldsymbol{Z}}(\boldsymbol{z})$$
$$\overline{\pi}^L_{R\&D} = \int_{\boldsymbol{z}} \frac{\mathbf{1}\{z_1 > k_L\} Pr^L(R\&D \mid Stay, \boldsymbol{z}, \boldsymbol{\theta})}{\eta^L} dG_{\boldsymbol{Z}}(\boldsymbol{z})$$
$$\overline{\pi}^S_{Exit} = \int_{\boldsymbol{z}} \frac{\mathbf{1}\{z_1 \leq k_L\} Pr^S(Exit \mid \boldsymbol{z}, \boldsymbol{\theta})}{1-\eta^L} dG_{\boldsymbol{Z}}(\boldsymbol{z})$$
$$\overline{\pi}^L_{Exit} = \int_{\boldsymbol{z}} \frac{\mathbf{1}\{z_1 > k_L\} Pr^L(Exit \mid \boldsymbol{z}, \boldsymbol{\theta})}{\eta^L} dG_{\boldsymbol{Z}}(\boldsymbol{z})$$
$$\overline{\pi}_{R\&D} = \int_{\boldsymbol{z}} \Big[\mathbf{1}\{z_1 \leq k_L\} Pr^S(R\&D \mid Stay, \boldsymbol{z}, \boldsymbol{\theta}) + \mathbf{1}\{z_1 > k_L\} Pr^L(R\&D \mid Stay, \boldsymbol{z}, \boldsymbol{\theta})\Big] dG_{\boldsymbol{Z}}(\boldsymbol{z})$$
$$\overline{\pi}_{Exit} = \int_{\boldsymbol{z}} \Big[\mathbf{1}\{z_1 \leq k_L\} Pr^S(Exit \mid \boldsymbol{z}, \boldsymbol{\theta}) + \mathbf{1}\{z_1 > k_L\} Pr^L(Exit \mid \boldsymbol{z}, \boldsymbol{\theta})\Big] dG_{\boldsymbol{Z}}(\boldsymbol{z})$$
$$\eta^L = \int_{\boldsymbol{z}} \mathbf{1}(z_1 > k_L) dG_{\boldsymbol{Z}}(\boldsymbol{z})$$

Now, suppose that the joint and marginal distributions of $(\varepsilon_{R\&D}, \varepsilon_{Exit})$ are continuous, so that all the resultant probabilities are also continuous (the distributions assumed in the R&D model satisfy this condition). Then, for an arbitrary value of the parameters $\boldsymbol{\theta}_2$ the right hand side of the equations presented above is a continuous function of the left hand side vector, $\boldsymbol{\theta}_1$. Therefore, the right hand side is a continuous mapping from $[0,1]^7 \times [0,1]^7$ and by Brower's Fixed Point Theorem, it has a fixed point. Since this is true for an arbitrary value of $\boldsymbol{\theta}_2$, it must hold for $\boldsymbol{\theta}_2^0$, the true values of these parameters. This proves that an equilibrium exists. □

Asymptotic properties of $\widehat{\boldsymbol{\theta}}^{\mathbf{EL}}$

Suppose the following conditions are satisfied:

- Ap.1 All equilibrium beliefs are strictly between 0 and 1.
- Ap.2 Identification conditions discussed in section 3.9.1 are satisfied.
- Ap.3 The log-likelihood $\log f(Y \mid \boldsymbol{Z}, \boldsymbol{\theta})$ satisfies the usual technical conditions for asymptotic consistency and normality of MLE.
- Ap.4 The sample jacobian matrix for equilibrium conditions $\frac{1}{N}\sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}} \Psi(\boldsymbol{z}_i, \boldsymbol{\theta})$ converges uniformly in probability to its expected value if $\boldsymbol{\theta}$ converges to $\boldsymbol{\theta}_0$.
- Ap.5 Technical conditions for the asymptotic normality of $\sqrt{N}\frac{1}{N}\sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}} \Psi(\boldsymbol{z}_i, \boldsymbol{\theta}_0)$ are satisfied

Let

$$\Im_0 = Var\big[\log f(Y \mid \boldsymbol{Z}, \boldsymbol{\theta}_0)\big], \quad A_0 = E\big[\nabla_{\boldsymbol{\theta}} \Psi(\boldsymbol{Z}, \boldsymbol{\theta}_0)\big], \quad B_0 = E\big[\Psi(\boldsymbol{Z}, \boldsymbol{\theta}_0) \Psi(\boldsymbol{Z}, \boldsymbol{\theta}_0)'\big]$$

Then, we have that:

$$\sqrt{N}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$$

$$\text{where} \quad \boldsymbol{\Omega} = \left(\Im_0 + A_0' B_0^{-1} A_0\right)^{-1}$$

Proof:

The corresponding Lagrangian for the EL estimation problem is given by:

$$\mathcal{L} = \sum_{i=1}^{N} \log f(y_i \mid \boldsymbol{z}_i; \boldsymbol{\theta}) + \sum_{i=1}^{N} \log p_i + \lambda(1 - \sum_{i=1}^{N} p_i) - N\boldsymbol{\nu}' \sum_{i=1}^{N} p_i \Psi(\boldsymbol{z}_i; \boldsymbol{\theta})$$

where $\lambda \in \mathbb{R}$ and $\boldsymbol{\nu} \in \mathbb{R}^7$ are Lagrange multipliers. First order conditions with respect to p_i yield:

$$\frac{1}{p_i} - \lambda - N\boldsymbol{\nu}'\Psi(\boldsymbol{z}_i; \boldsymbol{\theta}) = 0$$

multiplying both sides by p_i and summing over i yields

$$\lambda = N, \quad \text{and} \quad p_i = \frac{1}{N(1 + \boldsymbol{\nu}'\Psi(\boldsymbol{z}_i; \boldsymbol{\theta}))} \qquad \text{for } i = 1, .., N$$

Plugging back into the joint semi-empirical likelihood we get

$$\sum_{i=1}^{N} \log f(y_i \mid \boldsymbol{z}_i; \boldsymbol{\theta}) - \sum_{i=1}^{N} \log\left(1 + \boldsymbol{\nu}'\Psi(\boldsymbol{z}_i; \boldsymbol{\theta})\right) - N \log N$$

$\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}$ and $\boldsymbol{\nu}$ satisfy the first order conditions:

$$S_{1,N}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}, \boldsymbol{\nu}) \equiv \frac{1}{N} \sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}} \log f\left(y_i \mid \boldsymbol{z}_i; \widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\right) - \sum_{i=1}^{N} \frac{\nabla_{\boldsymbol{\theta}} \Psi\left(\boldsymbol{z}_i; \widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\right)' \boldsymbol{\nu}}{N\left\{1 + \boldsymbol{\nu}'\Psi\left(\boldsymbol{z}_i; \widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\right)\right\}} = 0$$

$$S_{2,N}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}, \boldsymbol{\nu}) \equiv \sum_{i=1}^{N} \frac{\Psi\left(\boldsymbol{z}_i, \widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\right)}{N\left\{1 + \boldsymbol{\nu}'\Psi\left(\boldsymbol{z}_i; \widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\right)\right\}} = 0$$

A first order Taylor series approximation around $(\boldsymbol{\theta}_0, \mathbf{0})$ yields

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} S_{1,N}^0 \\ S_{2,N}^0 \end{pmatrix} + \begin{pmatrix} I_N & -A_N' \\ A_N & -B_N \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\theta}}^{\mathrm{EL}} - \boldsymbol{\theta}_0 \\ \boldsymbol{\nu} \end{pmatrix} + o_p(N^{-1/2})$$

Where

$$S_{1,N}^0 = \frac{1}{N} \sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}} \log f(y_i \mid \boldsymbol{z}_i, \boldsymbol{\theta}_0), \quad S_{2,N}^0 = \sum_{i=1}^{N} \frac{1}{N} \Psi\left(\boldsymbol{z}_i, \boldsymbol{\theta}_0\right)$$

$$I_N = \frac{1}{N} \sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}, \boldsymbol{\theta}'} \log f(y_i \mid \boldsymbol{z}_i, \boldsymbol{\theta}_0), \quad A_N = \frac{1}{N} \sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}} \Psi(\boldsymbol{z}_i, \boldsymbol{\theta}_0), \quad B_N = \frac{1}{N} \sum_{i=1}^{N} \Psi(\boldsymbol{z}_i, \boldsymbol{\theta}_0) \Psi(\boldsymbol{z}_i, \boldsymbol{\theta}_0)'$$

If the assumptions Ap.1-Ap.5 are satisfied, then:

$$I_N \xrightarrow{p} \Im_0, \quad A_N \xrightarrow{p} A_0, \quad B_N \xrightarrow{p} B_0$$

and

$$\begin{matrix} \sqrt{N} S^0_{1,N} \\ \sqrt{N} S^0_{2,N} \end{matrix} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}), \quad \text{where} \quad \Sigma = \begin{pmatrix} \Im_0 & 0 \\ 0 & B_0 \end{pmatrix}$$

Therefore

$$\begin{pmatrix} \sqrt{N}(\widehat{\boldsymbol{\theta}}^{\mathbf{EL}} - \boldsymbol{\theta}_0) \\ \sqrt{N}\boldsymbol{\nu} \end{pmatrix} \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{\Omega}), \quad \text{where} \quad \Omega = \begin{pmatrix} -\Im_0 & -A_0' \\ A_0 & -B_0 \end{pmatrix}^{-1} \begin{pmatrix} \Im_0 & 0 \\ 0 & B_0 \end{pmatrix} \begin{pmatrix} -\Im_0 & -A_0' \\ A_0 & -B_0 \end{pmatrix}^{-1'}$$

and so we get

$$\sqrt{N}(\widehat{\boldsymbol{\theta}}^{\mathbf{EL}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}\big(\mathbf{0}, (\Im_0 + A_0' B_0^{-1} A_0)^{-1}\big)$$

as we claimed. □

Including additional information about $\boldsymbol{Z}$ increases efficiency of $\widehat{\boldsymbol{\theta}}^{\mathbf{EL}}$

Suppose assumptions Ap.1-Ap.5 are satisfied and we also know that the population distribution of $\boldsymbol{Z}$ satisfies:

$$E\big[\varphi(\boldsymbol{Z}, \boldsymbol{\xi}_0)\big] = 0$$

where $\varphi(\cdot) \in \mathbb{R}^r$ and $\boldsymbol{\xi}_0 \in \mathbb{R}^r$ is *known.*

Now suppose that we add the constraint $\sum_{i=1}^N p_i \varphi(\boldsymbol{z}_i, \boldsymbol{\xi}_0) = 0$ to the EL estimation. Let $\Im_0$, A_0 and B_0 be the same as in the previous proof and let

$$C_0 = E\big[\Psi(\boldsymbol{Z}, \boldsymbol{\theta}_0)\varphi(\boldsymbol{Z}, \boldsymbol{\xi}_0)'\big] \quad \text{and} \quad D_0 = E\big[\varphi(\boldsymbol{Z}, \boldsymbol{\xi}_0)\varphi(\boldsymbol{Z}, \boldsymbol{\xi}_0)'\big]$$

Then we have that

$$\sqrt{N}(\widehat{\boldsymbol{\theta}}^{\mathbf{EL}} - \boldsymbol{\theta}_0) \xrightarrow{d} \mathcal{N}\big(\mathbf{0}, \widetilde{\Omega}\big)$$

where

$$\widetilde{\Omega} = \big(\Im_0 + A_0' B_0^{-1} A_0 + A_0' B_0^{-1} C_0 \triangle_0^{-1} C_0' B_0^{-1} A_0\big)^{-1}, \quad \triangle_0 = D_0 - C_0' B_0^{-1} C_0$$

We have that $\widetilde{\boldsymbol{\Omega}}$ is smaller (in the positive semidefinite sense) than $\boldsymbol{\Omega}$ because the matrix $\triangle_0^{-1}$ is the lower sub-matrix of the inverse variance-covariance matrix of $\Psi(\boldsymbol{Z}, \boldsymbol{\theta}_0)$ and $\varphi(\boldsymbol{Z}, \boldsymbol{\xi}_0)$ and is therefore a positive definite matrix.

Proof:

Let $\boldsymbol{\nu}_{\boldsymbol{\varphi}}$ be the vector of Lagrange multipliers associated with the constraint $\sum_{i=1}^N p_i \varphi(\boldsymbol{z}_i, \boldsymbol{\xi}_0) =$

0. The first order conditions for the estimation problem are now

$$S_{1,N}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}},\boldsymbol{\nu},\boldsymbol{\nu}_{\boldsymbol{\varphi}}) \equiv \frac{1}{N}\sum_{i=1}^{N}\nabla_{\boldsymbol{\theta}}\log f\big(y_i \mid \boldsymbol{z}_i;\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\big) - \sum_{i=1}^{N}\frac{\nabla_{\boldsymbol{\theta}}\Psi\big(\boldsymbol{z}_i;\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\big)'\boldsymbol{\nu}}{N\big\{1+\boldsymbol{\nu}'\Psi\big(\boldsymbol{z}_i;\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\big)+\boldsymbol{\nu}_{\boldsymbol{\varphi}}'\varphi(\boldsymbol{z}_i,\boldsymbol{\xi}_0)\big\}} = 0$$

$$S_{2,N}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}},\boldsymbol{\nu},\boldsymbol{\nu}_{\boldsymbol{\varphi}}) \equiv \sum_{i=1}^{N}\frac{\Psi\big(\boldsymbol{z}_i,\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\big)}{N\big\{1+\boldsymbol{\nu}'\Psi\big(\boldsymbol{z}_i;\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\big)+\boldsymbol{\nu}_{\boldsymbol{\varphi}}'\varphi(\boldsymbol{z}_i,\boldsymbol{\xi}_0)\big\}} = 0$$

$$S_{3,N}(\widehat{\boldsymbol{\theta}}^{\mathrm{EL}},\boldsymbol{\nu},\boldsymbol{\nu}_{\boldsymbol{\varphi}}) \equiv \sum_{i=1}^{N}\frac{\varphi\big(\boldsymbol{z}_i,\boldsymbol{\xi}_0\big)}{N\big\{1+\boldsymbol{\nu}'\Psi\big(\boldsymbol{z}_i;\widehat{\boldsymbol{\theta}}^{\mathrm{EL}}\big)+\boldsymbol{\nu}_{\boldsymbol{\varphi}}'\varphi(\boldsymbol{z}_i,\boldsymbol{\xi}_0)\big\}} = 0$$

The result follows from a first-order Taylor series approximation around the point $(\boldsymbol{\theta}_0,\mathbf{0},\mathbf{0})$.

Appendix 2: Estimation results

TABLE 1
Empirical Likelihood estimates for strategic coefficients
(Standard errors in parentheses.)

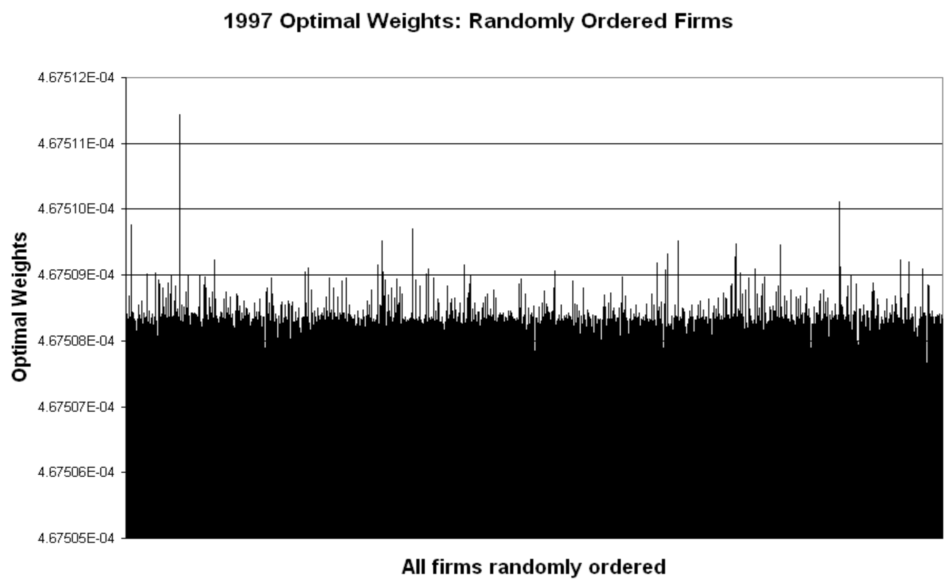
	1997	1998	1999
$\widetilde{\alpha}_1$	4.4974	2.4887	5.3535
	(1.2318)	(0.7654)	(1.2590)
$\widetilde{\alpha}_2$	−4.4539	−3.20197	−5.6884
	(1.3406)	(0.9319)	(1.3868)
$\widetilde{\gamma}_1$	32.8797	−27.7714	−16.8313
	(129.7842)	(12.3511)	(4.3046)
$\widetilde{\gamma}_2$	-4.0045	60.0549	38.4981
	(129.6698)	(14.3920)	(7.2893)
α_1^S	2.4630	1.1448	2.7582
	(0.6452)	(0.4503)	(0.6651)
α_2^S	−2.4195	−1.8582	−3.0931
	(0.7297)	(0.5420)	(0.7570)
γ_1^S	31.0506	-2.5684	0.7332
	(70.6443)	(7.1888)	(2.4752)
γ_2^S	-2.1754	34.8519	20.9335
	(70.4424)	(8.3777)	(3.9836)
α_1^L	2.0778	0.6304	2.2604
	(0.5400)	(0.3685)	(0.5593)
α_2^L	−2.0343	−1.3438	−2.5953
	(0.6143)	(0.3934)	(0.6369)
γ_1^L	30.7043	7.0805	4.1020
	(59.4655)	(5.6465)	(2.4713)
γ_2^L	-1.8291	25.2030	17.5645
	(59.2274)	(6.0945)	(3.3668)
σ	-0.6583	-0.6595	-0.67488

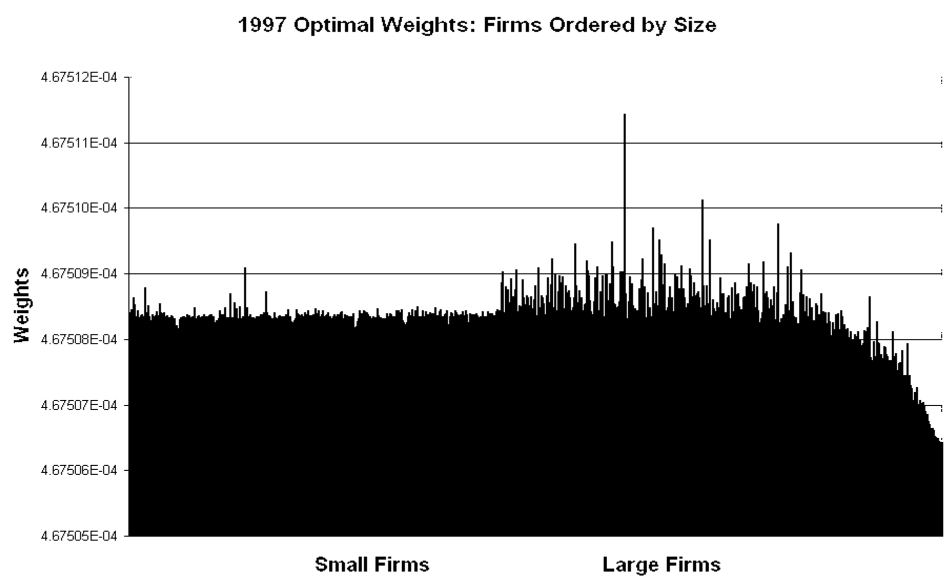
TABLE 2
Empirical likelihood estimates for beliefs
(Standard errors in parentheses.)

	1997	**1998**	**1999**
$\overline{\pi}^L_{R\&D}$	0.6569	0.5755	0.6392
	(0.0145)	(0.0173)	(0.0168)
$\overline{\pi}^S_{R\&D}$	0.6174	0.4778	0.5891
	(0.0162)	(0.0203)	(0.0185)
$\overline{\pi}^L_{Exit}$	0.0725	0.1027	0.0967
	(0.0076)	(0.0101)	(0.0098)
$\overline{\pi}^S_{Exit}$	0.0767	0.0829	0.0741
	(0.0085)	(0.0133)	(0.0095)
η^L	0.5432	0.5603	0.5437
	(0.0107)	(0.0125)	(0.0122)

TABLE 3
Empirical likelihood estimates for private information variables
(Standard errors in parentheses.)

		R&D			**Exit**	
	1997	**1998**	**1999**	**1997**	**1998**	**1999**
q	0.0042	0.0038	0.0010	0.0014	0.00721	0.0039
	(0.0031)	(0.0027)	(0.0011)	(0.0031)	(0.0035)	(0.0017)
eps	0.0332	0.2989	0.1212	0.04852	0.1189	0.0375
	(0.0211)	(0.0452)	(0.0361)	(0.0247)	(0.0514)	(0.0506)
$\frac{\text{Empl}}{\text{Avg.Empl}}$	0.9296	−0.5382	-0.0114	3.5022	1.5967	2.3501
	(0.3982)	(0.4546)	(0.3791)	(1.6158)	(1.2578)	(1.1688)
Techsegment	0.17206	0.1649	0.2369	0.0855	−0.2435	0.1392
	(0.0493)	(0.0601)	(0.0571)	(0.0848)	(0.1103)	(0.0670)
$\frac{\text{Intangible assets}}{\text{Total assets}}$	-0.6498	-0.1973	-0.8397	NA	NA	NA
	(0.4063)	(0.4380)	(0.3896)	NA	NA	NA
$\frac{\text{Long term debt}}{\text{Total assets}}$	NA	NA	NA	-0.1688	0.0862	−0.5210
	NA	NA	NA	(0.3246)	(0.3487)	(0.2061)


1997 Optimal Weights: Randomly Ordered Firms
4.67512E-04
4.67511E-04
4.67510E-04
4.67509E-04
4.67508E-04
4.67507E-04
4.67506E-04
4.67505E-04
Optimal Weights
All firms randomly ordered


1997 Optimal Weights: Firms Ordered by Size
4.67512E-04
4.67511E-04
4.67510E-04
4.67509E-04
4.67508E-04
4.67507E-04
4.67506E-04
4.67505E-04
Weights
Small Firms
Large Firms

References

Brock, W.A and Durlauf ,S.N (2001) 'Interactions-based models.' *In* Heckman, J.J and Leamer, E, eds.*Handbook of Econometrics*, Vol. V, Chapter 54 Elsevier Science B.V, Amsterdam.

Cosslett, S.R. (1981). 'Maximum likelihood estimator for choice-based samples' *Econometrica* **49**, 1289-1316.

Hall and Vopel (1997). 'Innovation, market share and market value' NBER.

Hall (1999). 'Innovation and market value' NBER Working Paper No. 6984

Hellerstein, J.K. and Imbens, G.W. (1999). 'Imposing moment restrictions from auxiliary data by weighting' *The Review of Economics and Statistics* **81** 1-14.

Imbens, G.W. (1992). 'An efficient method of moments estimator for discrete choice models with choice-based sampling' *Econometrica* **60** 1187-1214.

Imbens, G.W. and Lancaster, T. (1994) 'Combining micro and macro data in microeconometric models' *The Review of Economic Studies* **61** 655-680.

Jovanovic, B and MacDonald, G.M. (1994) 'The life cycle of a competitive industry.' *Journal of Political Economy* **102** 322-347.

Kitamura, Y. (2001). 'Asymptotic optimality of empirical likelihood for testing moment restrictions.' *Econometrica* **69** 1661-1672.

Kleppler, S. (1996). 'Entry, exit, growth and innovation over the product life cycle.' *American Economic Review* **86** 562-583.

Klepper, S. and Simons, K. (2000) 'The making of an oligopoly: Firm survival and technological change in the evolution of the U.S tire industry' *Journal of Political Economy* **102(2)** 728-760.

Lazar, N.A. (2000). 'Bayesian empirical likelihood' Technical report, Carnegie Mellon University, Department of Statistics.

Lieberman, M.V (1990). 'Exit from declining industries: shakeout or stakeout?' *Rand Journal of Economics* **21** 438-454.

Nelsen, R.B (1999). *An Introduction to Copulas.* Springer-Verlag: New York.

Newey, W.K and Smith, R.J (2001). 'Higher order properties of GMM and generalized empirical likelihood estimators.' Mimeo

Oriani, R. and Sobrero, M. (2003). 'Assessing the market valuation of firms' technological knowledge using a real options perspective'. Mimeo

Owen, A.B. (1988). 'Empirical likelihood ratio confidence intervals for a single functional' *Biometrika* **75**, 237-249.

Owen, A.B. (1990). 'Empirical likelihood confidence regions' *The Annals of Statistics* **18**, 90-120.

Owen, A.B. (1991). 'Empirical likelihood for linear models' *The Annals of Statistics* **19**, 1725-1747.

Owen, A.B. (2001). *Empirical Likelihood.* Monographs on Statistics and Applied Probability 92. Chapman & Hall/CRC.

Qin, J. (1994). 'Semi-empirical likelihood ratio confidence intervals for the difference of two sample means' *Annals of the Institute of Mathematical Statistics* **46**, 117-126.

Qin, J. (1998). 'Semiparametric likelihood based method for goodness of fit tests and estimation in upgraded mixture models' *Scandinavian Journal of Statistics* **25** 681-691.

Qin, J. (2000). 'Combining parametric and empirical likelihoods' *Biometrika* **87**, 484-490.

Reinganum, J.F (1989) 'The timing of innovation: research, development and diffusion.' *In* Schmalensee, R. and Willig, R., eds.*Handbook of Industrial Organization*, Vol. I, Chapter 14 Elsevier Science B.V, Amsterdam.

Seim, K. (2002). 'An empirical model of firm entry with endogenous product-type choices.' Mimeo, Graduate School of Business, Stanford University, Palo Alto, CA.

Tamer, E. (2003). 'Incomplete simultaneous discrete response model with multiple equilibria.' *Review of Economic Studies* **70**, 147-165.