EL COLEGIO DE MÉXICO, A.C.
CENTRO DE ESTUDIOS ECONÓMICOS

"ECONOMETRIC ESTIMATION OF STRATEGIC INTERACTION MODELS"

TESIS PRESENTADA POR:

BRASIL ALBERTO ACOSTA PEÑA

PARA OPTAR POR EL GRADO DE

DOCTOR EN ECONOMÍA

PROMOCIÓN AÑO 2005-2008

DIRECTOR DE TESIS

ANDRÉS ARADILLAS LÓPEZ

MÉXICO D.F. AGOSTO DE 2009

EL COLEGIO DE MÉXICO

CENTRO DE ESTUDIOS ECONÓMICOS

CONSTANCIA DE APROBACIÓN

Director de Tesis: Andrés Aradillas López

Aprobada por el Jurado Examinador:

Dr. Angel Calderón Madrid, Presidente ________________________________

Dr. Edwin Van Gameren, Primer Vocal ________________________________

Dr. Eneas Arturo Caldiño García,
Vocal Secretario ________________________________

Dr. Ignacio N. Lobato, Lector externo ________________________________

Dr. Carlos Chiapa Labastida, Suplente ________________________________

Table of Contents

Introduction

Econometric analysis of qualitative response models was developed by McFadden (1981). The dependent variable included in the analysis is intrinsically categorical. Qualitative variables could be binomial (yes/no), or multinomial. In the multinomial case, models can be ordered or unordered. In ordered models, as the name suggests, the dependent variable will be an ordered response; for example, might be a credit rating on a scale from zero to six. This variable will assume 0, in the lowest category, and 6 in the greatest one, and we try to estimate the probability of each choice. In unordered models there is no particular restriction in the values of the dependent variable, but this variable still being discrete (see Wooldrige, 2002 for details). Then, we can assume that agents should make their choices in a qualitative sense. But this models do not capture the interaction between the agents. Interaction means interdependence between individual decisions which are not mediated by markets, Brock and Durlauf (2001); they study "interaction-based models", which are mathematically equivalent to logistic models of discrete choice. By interaction-based models, they refer to "a class of economics environment in which the payoff function of a given agent takes as direct arguments the choices of the other agents".

A natural extension of interaction-based includes *asymmetric information*, using game theoretical foundations. Indeed, Aradillas-Lopez (2003), proposed an estimator in order to find the interaction coefficients in a context of economic interaction models with asymmetric information. "The presence of *asymmetric information* implies that agents must construct beliefs about other agents". If we assume that observed choices are derived from a Bayesian-Nash equilibrium, these beliefs must be rational and satisfy the conditions consistent with such an equilibrium.

Chapter 1: Econometric models under complete information environment with game theoretic foundations are used in this chapter to analyze husband and wife labor force participation. It is allowed multiple equilibria and was designed a mechanism of equilibrium's selection. Following Aradillas-López (2008), is used a consistent estimator of interaction parameters α's in the context of non unique Nash-Equilibrium. We used "Basal Survey of Savings, Credit and Rural Micro-Finances" in Mexico, made by BANSEFI as data set in 2004. And it was found that married couples interact strategically. In average, wives decide not to participate in the labor market if there husbands participate. A brief survey of games and econometrics is included.

Chapter 2: A semi-empirical likelihood estimator is proposed for models where agents interact under asymmetric information. The methodology focuses on situations where some variables that were privately observed when choices were made become available to the econometrician afterwards. This variables are assumed to have a finite support. The main feature of the estimator is that structural parameters, beliefs and unknown probability distribution function of these privately observed variables are estimated simultaneously under the assumption that observed outcomes are the result of a Bayesian-Nash Equilibrium. The methodology is applied to three actions and three types of agents. Firms decide to be aggressive, neutral or passive in their investment decisions. Estimation shows a significant component of strategic interaction in the case of small and medium size (type) of firms. Interaction is more significant to small firms than the others.

Chapter 3: This chapter analyzes the technical efficiency in the Mexican manufacturing sector in which determinants and changes of the efficiency since NAFTA

(North American Free Trade Agreement) are studied using the Panel Data Stochastic Frontier Analysis in its time-invariant and time-variant versions, comparing each other. It was used the Annual Industry Survey (AIS), which panel data information allow us to model the efficiency performance of firms in the period 1994-2001. Our main findings show that Mexican manufacturing firms worked, in average, at almost 23% of its potential product (compared with the best firm performance) and that there was a slight lost of capacity along time (1994-2001). Additionally, could be detected structural change, understood it as the change of firms' ranking observed in the model and the coefficient of the production function. Moreover, could be detected those firms that were consistent, winners or losers in the process of openness. Finally, under the assumption of Cournot duopoly competition, there are studied determinants for R&D investment.

Chapter 1

Estimation of a Multiple Equilibrium Game with Complete Information: Husband and Wife Labor Force Participation in Mexico

1.1 Introduction

Labor participation has been studied in different contexts, using different methodologies. Gong and Soest (2002), for example, investigate labor supply of married women in Mexico City, using a neoclassical structural model. They found that income elasticity of labor supply is about -0.17, and wage elasticity, is about 0.18. Other way to focus the same problem has been studied under the named *gender problem*. Greenstein (1996), studied "the interaction between the ideologies of wives and their husbands in order to understand how a division of household labor emerges", "husbands do relatively little domestic labor unless both they and their wives are relatively egalitarian in their beliefs about gender and marital roles." In terms of labor supply, husband could have influence over his wife in a negative way, in order to forbid her to participate in labor market.

Calderon-Madrid (2007), using survival analysis, studied "determinants of time spent in formal and informal sectors by employees, as well as to assess the costs for workers resulting from the mobility between formal and informal job status"; he found that "workers with less human capital are more likely to be trapped in cycles of long spells of informal employment followed by short-term jobs in the formal sector and displaced again to informal jobs". Orley Ashenfelter (1984), studied the relationship between macroeconomic analysis and microeconomic analysis of labor supply.

On the other hand, in the context of game theoretical foundations, McFadden (1974) developed single-person qualitative choice model, and proposed the well known *probit* and *logit* model (see Wooldrige, 2002, pg. 457-516). Heckman (1978), studied models with structural shift parameters. Brock and Durlaf (2001), study "interaction-based models". They define interaction as interdependencies between individual decisions which are not mediated by markets between the agents. By interaction-based models, they refer to "a class of economics environment in which the payoff function of a given agent takes as direct arguments the choices of the other agents". Bjorn and Vuong (1984), used a game theoretic approach formulating simultaneous equations models for a dummy endogenous variables applied to a study of husband and wife labor force participation. More over, Bjorn and Vuong (1985) extend their pioneer work introducing Stackelberg equilibrium in which husband is the leader, i.e., the husband knew what action his wife would take and he optimizes accordingly.

Kooreman (1994), used Bjorn and Vuong's Nash and Stackelberg econometric framework (likelihood estimation), introducing imposed Pareto optimality. He studies joint labor force participation decisions of husbands and wives in a sample of Dutch households. In Stackelberg game, he allows husband and wife to be leader respectively.

Nonetheless, under this framework, economists "have made simplifying assumptions to avoid multiplicity"[1] of equilibrium. Tamer (2003), introduced a bivariate simultaneous discrete response model "which is a stochastic representation of equilibrium

[1] This is the case of Bjorn and Vuong (1984, 1985) and Kooreman (1994), among others.

in a two-person discrete game". He analyzed the model in the presence of *multiple equilibrium* and proposed a parametric and non parametric estimator avoiding to invoke the coherency condition (see bellow) or imposing *ad-hoc* structure of the game in order to fit it in a well behave environment. Pakes, Ostrovsky, and Berry (2004) study the players' strategies within the structure of discrete, dynamics games. Their model is constructed within a basic dynamic game of entry and exit. Golan, Karp, and Perloff (2007) have a Chapter in their book that study how to estimate strategies using generalized maximum entropy (GME) framework. Bresnahan and Reiss (1991), extended qualitative choice models developing econometric models for discrete games. They modeled the payoff of games where an econometrist observes qualitative or censored information about agents' decisions and payoffs. Equations describing players' equilibrium strategies depend on the game's structure and the equilibrium solution concept. They showed that one can describe the equilibria of a simultaneous-move Nash game with a linear system of dummy endogenous variables. They also showed that sequential-move and cooperative models have different, but related, econometric structure. They apply this framework to models of market entry, technology adoption, tax auditing, and *cooperative family labor supply*. Under incomplete informational game-theoretical environment, Aradillas-Lopez (2003, 2008) has proposed likelihood-based procedures "where players unobserved beliefs and the vector of payoff parameters are estimated simultaneously by solving a well defined sample analog of the population equilibrium conditions". He has proposed, for example, a semi-empirical likelihood estimator where agents interact under asymmetric information. He carefully study uniqueness and existence equilibrium problems. Nonetheless, it is difficult to impalement this kind of models because of lack information in labor participation surveys in Mexico: they do not capture the incomplete information environment.

Here, I studied husband and wife labor force participation in Mexico, using estimation of a simultaneous game with *complete information*, allowing multiplicity of equilibria, including a selection mechanism of the optimal equilibrium. Main questions are two:

a married couple interact strategically when they make their participation decisions to enter or not in the labor market in Mexico?, or exists some sociocultural items that lead decisions of participation in the labor market of married people in Mexico?

1.2 Games and econometrics

A strategic game is a model of interacting decision-makers: *players*. Each player has a set of possible *actions*. The model captures interaction between the players by allowing each player to be affected by the actions of *all* players. Each player has *preferences* about the action *profile* (the list of all players' actions). Following Osborne (2004, pg. 13-14) a *strategic game* (with ordinal preferences) consist of

1. a set of *players*
2. for each player, a set of *actions*
3. for each player, *preferences* over the set of actions profiles.

Can the agents arrive to an equilibrium state given the interaction between them? The answer was given by John Nash. Assuming that the players are rational, i.e., they choose the best available action; it is possible to reach an equilibrium (no player has incentives to change the status in which each player is in equilibrium). Nash equilibrium of strategic game with ordinal preferences is defined as follows:

> **Definition:** The action profile a^* in a strategic game with ordinal preferences is a **Nash equilibrium**, if for every player i and every action a_i of a player i, a^* is at least as good according to player *i's* preferences as the action profile (a_i, a^*_{-i}) in which player i chooses a_i while every other player j chooses a^*_j. Equivalently, for every player i,
>
> $$u_i(a^*) \geq u_i(a_i, a^*_{-i})$$

for every action a_i of player i, where u_i is a payoff function that represents player *i's* preferences[2]. Osborne (2004, pg. 23).

A generalization of Nash equilibrium can be made. We allow each player to choose a probability distribution over his set of actions rather than restricting him to choose deterministic action. Then, *mixed strategy* can be defined as a probability distribution over the player's action. It is important to notice that a mixed strategy may assign probability 1 to a single action: by allowing a player to choose probability distributions, we do not prohibit to the players from choose deterministic actions. This kind of "mixed strategy" can be considered as a *pure strategy*, Osborne (2004, pg. 107-108).

There is another class of environment in which agents can interact strategically: *complete information* or *asymmetric information* context. Asymmetric means that some parties are informed about variables that affect everyone, and some parties are not.

1.2.1 Discrete Strategy Game

Complete Information

Let be a simultaneous 2×2 game in its normal representation.

Figure 1

A simple 2 x 2 game

		PLAYER 2	
		Y=1	Y=0
PLAYER 1	Y=1	$t_1 + \alpha_1, t_2 + \alpha_2$	$t_1, 0$
	Y=0	$0, t_2$	0,0

[2] a_{-i}, means actions of players different to player i.

Where, in general, each player has two mutually exclusive actions: Y=0 or Y=1 (participate or don't; to be aggressive or don't; enter or don't, etc.). Players' payoffs depend on their actions: players will receive $(t_1 + \alpha_1, t_2 + \alpha_2)$ if they choose $(Y = 1, Y = 1)$; $(0, 0)$ if they choose $(Y = 0, Y = 0)$; $(0, t_2)$ if they choose $(Y = 0, Y = 1)$, and $(t_1, 0)$ if they choose $(Y = 1, Y = 0)$. Notice that (α_1, α_2) appear only in the case in which each player choose Y=1, separately or together. $\alpha_p' s$ try to capture how other player's action affects player p, for $p = \{1, 2\}$; they are known as "interaction coefficients".

Regardless the signs of alpha's, based on the game in the Figure 1, it is true that:

If $t_1 + \alpha_1 \geq 0$ and $t_2 + \alpha_2 \geq 0$ players will choose (1,1)

If $t_1 < 0$ and $t_2 < 0$ players will choose (0,0)

If $t_1 \geq 0$ and $t_2 + \alpha_2 < 0$ players will choose (1,0)

If $t_1 + \alpha_1 < 0$ and $t_2 \geq 0$ players will choose (0,1)

Now, following McFadden (1974), payoffs can be treated as random decomposed into deterministic components and random components. Let $(X_1, \varepsilon_1) \in \mathbb{R}^k \times \mathbb{R}$ and $(X_2, \varepsilon_2) \in \mathbb{R}^k \times \mathbb{R}$. Assume that $(\varepsilon_1, \varepsilon_2) \simeq F(\cdot; \Omega)$, where $F(\cdot)$ is known and Ω (variance and covariance matrix), unknown. Let be:

$$t_1 \equiv X_1' \beta_1 - \varepsilon_1$$
$$t_2 \equiv X_2' \beta_2 - \varepsilon_2$$

Where $\mathbf{X} \equiv (X_1, X_2)$ is the characteristics vector[3]; $\boldsymbol{\beta} \equiv (\beta_1, \beta_2)$, unknown parameters (deterministic part), and $\boldsymbol{\varepsilon} \equiv (\varepsilon_1, \varepsilon_2)$, unknown (for the econometrician) error term (random part). If only *pure strategies* are considered, players' optimal actions are simply given by

$$Y_p = 1\{X_p' \beta_p + \alpha_p Y_{-p} - \varepsilon_p \geq 0\} \tag{1.1}$$

[3]This variables will depend on the nature of the specific game.

For $p = 1, 2$. Where Y_{-p} means the action taken by the p player's opponent, and $1\{A\}$ is the indicator function: $1\{A\} = 1$ if A is true, 0 otherwise. Here the objective is estimate the parameters:

$$\boldsymbol{\theta} = (\beta_1, \beta_2, \alpha_1, \alpha_2, \Omega)$$

This is a well known model in econometrics studied by Heckman (1978), Schmidt (1982) and many others. The key issue is *Statistical Coherency* which is a necessary and sufficient condition for the likelihood of the model to be well defined:

$$Pr[(1,1)|\mathbf{X}] + Pr[(0,0)|\mathbf{X}] + Pr[(0,1)|\mathbf{X}] + Pr[(1,0)|\mathbf{X}] = 1$$

$\Leftrightarrow \alpha_1 \times \alpha_2 = 0$. But this *coherency condition* especially eliminates simultaneity from the model (Kooreman, 1994).

Now assuming that the *econometrician knows the signs* of the $\alpha's$, and that following assumptions hold:

Assumption A1 (**Information structure**)

i Realizations of (X_p, ε_p) are common knowledge. There is no sources of private information.

ii More over, in the context of *complete information*[4] each player knows their opponent action (or strategy) Y_{-p}.

Assumption A2 (**Strategic behavior**)

i Under *complete informational* assumption, players could play pure or mixed strategies, then multiple equilibria can be allowed.

[4]When each player doesn't know its opponent action (strategy), we say there is *incomplete information* and players should construct their beliefs about the other player action. Aradillas-Lopez (2008).

Assumption A3 (**Distributional properties of ε_1, ε_2, and η**)

i $(\varepsilon_1, \varepsilon_2)$ are jointly continuously distributed random variables with unbounded support. They are allowed to be correlated, but are assumed to be independent of all other variables in the model, known as orthogonality condition. The conditional support $\mathbb{S}(\varepsilon_p|\varepsilon_{-p})$, is assumed to be unbounded for $p = 1, 2$, for any possible realization of ε_{-p}

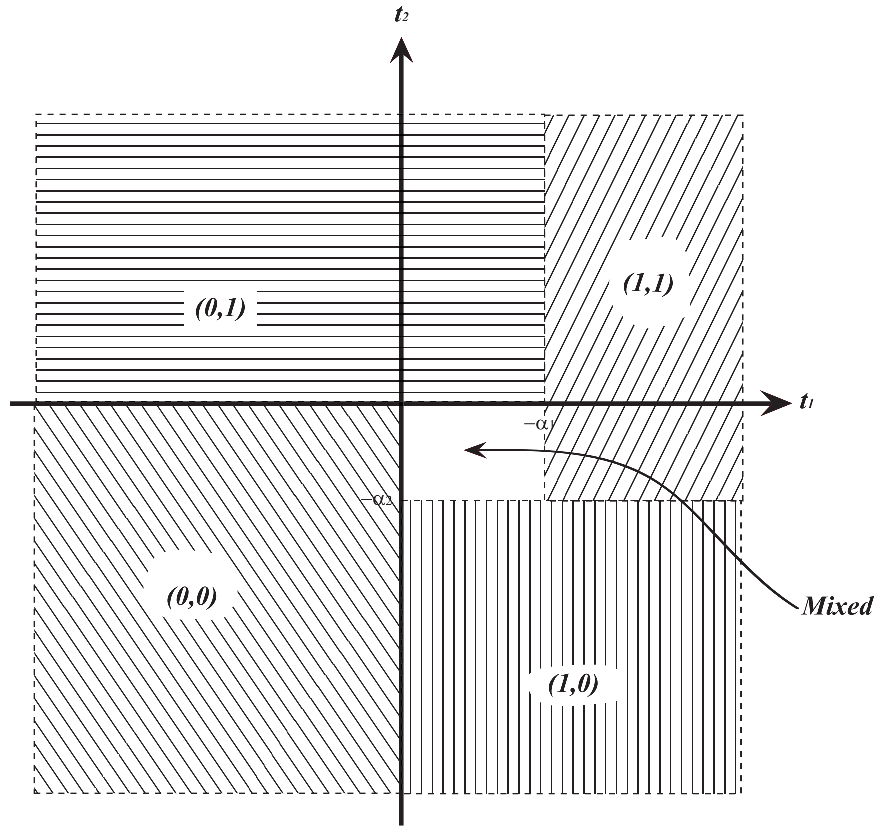
ii $G_p(\epsilon_p)$ will denote the marginal distribution or ε_p, with density $g_p(\epsilon_p)$. The joint distribution of $(\varepsilon_1, \varepsilon_2)$ is given by $G_{1,2}(\varepsilon_1, \varepsilon_2; \rho)$, where $\rho \in [-1, 1]$ summarize the entire dependence between ε_1 and ε_2; $G_{1,2}(\varepsilon_1, \varepsilon_2; \rho)$ represents unobserved (to the econometrician) distribution profits. For a given value of ϵ_1 and ϵ_2, the joint distribution of $G_{1,2}$ is an invertible function of ρ. This is true for all $(\epsilon_1, \epsilon_2) \in \mathbb{R}^2$.

Pure Equilibria

Now, assuming that the econometrist knows that $\alpha_1 \leq 0$, and $\alpha_2 \geq 0$ (parallel results arrive when $\alpha_1 \geq 0$, and $\alpha_2 \leq 0$), R^2 would be partitioned as follows:

Figure 2

(R^2 **Partition Pure Equilibrium**)



In the blank square named "Mixed" either outcome is likely. Here there are only pure equilibria but is not clear how to proceed in this case.

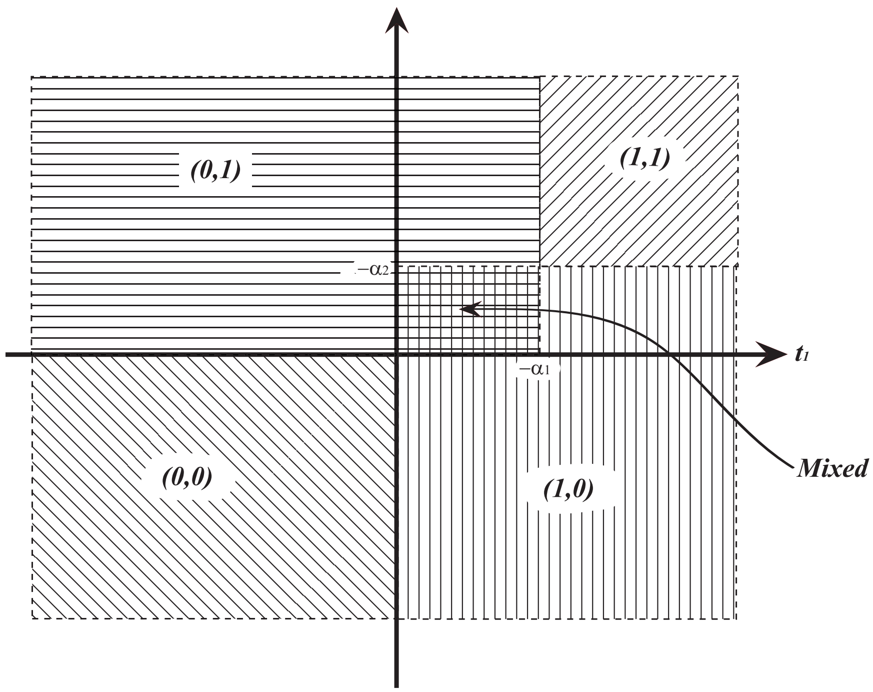
Multiple Equilibria

Now, under the assumption that the econometrist knows that $\alpha_1 \leq 0$, and $\alpha_2 \leq 0$ (when $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$, results are very similar). Let be $\mathbb{S}(M)$ the support of the random variable M, then, if $\mathbb{S}(\varepsilon_1, \varepsilon_2) \in \mathbb{R}^2$, the game with *complete information* will have *multiple equilibria* with positive probability for any realization of $\mathbf{X}$ unless: (a) $\alpha_1 \times \alpha_2 \leq 0$, if mixed-strategies are allowed (previous case), or (b) $\alpha_1 \times \alpha_2 = 0$, if mixed-strategies are ruled out (*coherency condition*). Now, we can do a partition of

R^2, drawing the regions conformed by the solution of the game when $\alpha_1 \leq 0$, and $\alpha_2 \leq 0$, then:

Figure 3

(R^2 **Partition Multiple Equilibrium**)



Then we have five regions. Pure strategy equilibrium, $\mathbf{R}_{(1,1)}$, $\mathbf{R}_{(0,0)}$, $\mathbf{R}_{(0,1)}$, and $\mathbf{R}_{(1,0)}$; mixed strategy equilibrium, $\mathbf{R}_{square}$.

In the middle box ($\mathbf{R}_{square}$) there are multiple Nash-equilibrium: two pure, (0,1) and (1,0), and one mixed. By definition[5], mixed strategy equilibrium means that player 2 will choose $Y_2 = 1$ with probability Π_2, and player 1 will choose $Y_1 = 1$ with probability Π_1. But player 2 chooses Π_2 such that player 1 is indifferent between $Y_1 = 1$ and $Y_1 = 0$; player 1, analogously, will choose Π_1 such that player 2 would be indifferent between $Y_2 = 1$ and $Y_2 = 0$. Equalizing expected utility of $Y_1 = 1$ and $Y_1 = 0$ we found that:

$$\Pi_2 = -\tfrac{t_1}{\alpha_1}, \text{ and } \Pi_1 = -\tfrac{t_2}{\alpha_2}$$

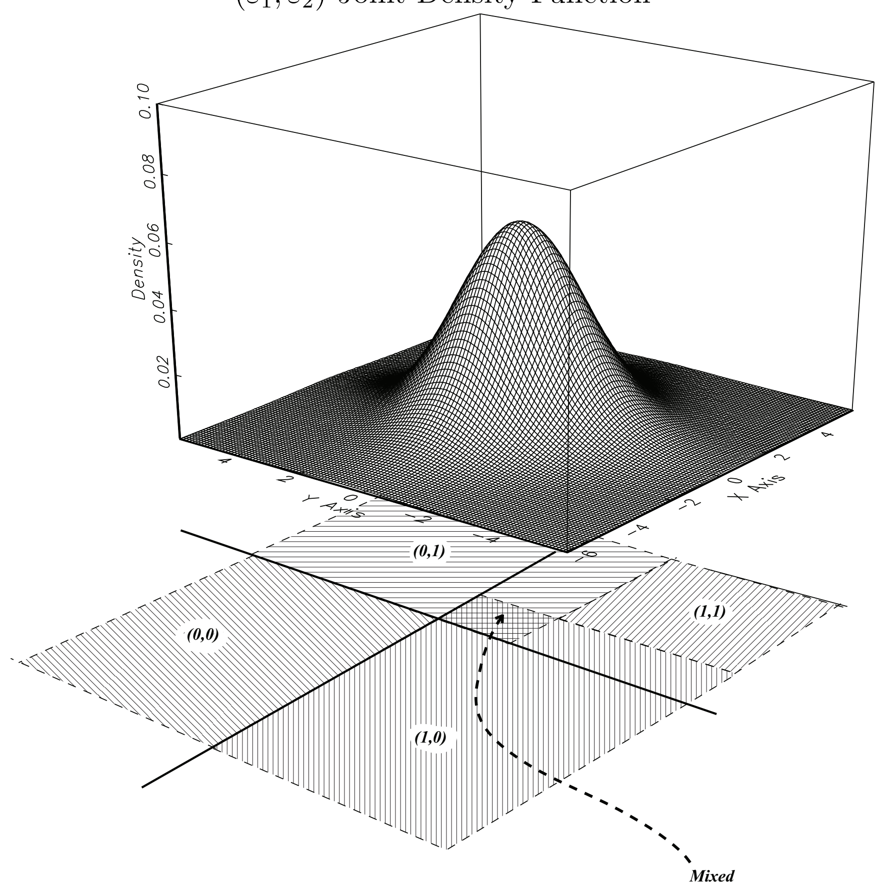
[5]See Osborne (pg. 137-142).

where $(\Pi_1, \Pi_2) \in [0,1] \times [0,1]$, $\forall$ t_1, t_2, α_1 and α_2 in the multiple equilibrium region ($\mathbf{R}_{square}$).

In this framework it is allow statistical interdependence between ε_1 and ε_2. Then, each region will aport certain amount of probability, according to the joint distribution of ε_1 and ε_2. This can be seen in Figure 4, in which is assumed that $(\varepsilon_1, \varepsilon_2)$ follows a Farlie-Gumbel-Morgesten families of joint distributions (see Johnson, et. al. 1999).

Figure 4

$(\varepsilon_1, \varepsilon_2)$ Joint Density Function



The main problem here is that:

$$Pr[(1,1)|\mathbf{X}] + Pr[(0,0)|\mathbf{X}] + Pr[(0,1)|\mathbf{X}] + Pr[(1,0)|\mathbf{X}] > 1$$

Econometric models of some discrete games

We are going to survey briefly some econometrics models of discrete games. Pioneer work is Bjorn and Vuong (1984,1985). They fit into the econometric model the interaction between two players (wife and husband), under the assumption that the outcomes of the database came from a Nash and Stackelberg equilibrium. Kooreman (1994), go further and proposed Pareto optimality. Finally, Bresnahan and Reiss (1991) and Tamer (2003) elegantly showed, identification of $\boldsymbol{\theta}$ is still viable in the *complete information* setting if there exists prior knowledge of the signs of strategic interaction, $\alpha's$. Following Kooreman (1994), general setting of Nash equilibrium (NE), Stackelberg equilibrium (SE) and Pareto optimality (PO) can be summary as follows:

1. There are *2 players*: (Wife and Husband; Small Firm and Large Firm, etc.)

2. There are *2 actions*. The action of player p is represented by a dummy variable for $p = \{1, 2\}$:

$$Y_p = \begin{cases} 1, & \text{enter, be aggressive, participate, etc.;} \\ 0, & \text{don't enter, don't be aggressive, don't participate, etc.} \end{cases}$$

3. *Preferences* are constructed as follows. Let be

$$U_p(Y_1, Y_2)$$

utility function for player p. Combination of action can be named *allocation*. Assuming ordinal preferences, each player can ranking its four utility levels. This imply that there is 4! different rankings which means that for two players there are $(4!)^2 = 576$ possible combination of utility ranks. For empirical implementation let be:

$$\begin{array}{ll}
\textbf{Player 1} & \textbf{Player 2} \\
U_1(1,1) = X'\beta_1^1 + \alpha_1^1 + \varepsilon_1^1 & U_2(1,1) = X'\beta_1^2 + \alpha_1^2 + \varepsilon_1^2 \\
U_1(1,0) = X'\beta_1^1 \qquad\;\; + \varepsilon_1^1 & U_2(1,0) = X'\beta_1^2 \qquad\;\; + \varepsilon_1^2 \\
U_1(0,1) = X'\beta_0^1 + \alpha_0^1 + \varepsilon_0^1 & U_2(0,1) = X'\beta_0^2 + \alpha_0^2 + \varepsilon_0^2 \\
U_1(0,0) = X'\beta_0^1 \qquad\;\; + \varepsilon_0^1 & U_2(0,0) = X'\beta_0^2 \qquad\;\; + \varepsilon_0^2
\end{array} \tag{1.2}$$

As mentioned, it was used McFadden's random utility hypothesis. Notice that there are eight parameters to be identified:

$$\{\beta_0^1, \beta_1^1, \beta_0^2, \beta_1^2, \alpha_0^1, \alpha_1^1, \alpha_0^2, \alpha_1^2\}$$

Which is difficult to be solved. Then, an alternative representation of (1.2) can be proposed:

$$\begin{array}{ll}
\textbf{Player 1} & \textbf{Player 2} \\
U_1(1,1) - U_1(1,0) = \alpha_1^1 & U_2(1,1) - U_2(1,0) = \alpha_1^2 \\
U_1(0,1) - U_1(0,0) = \alpha_0^1 & U_2(0,1) - U_2(0,0) = \alpha_0^2
\end{array} \tag{1.3}$$

This representation reduces the number of possible utility rankings per player from 24 to 6. For example, if econometrist knows that $\alpha_1^1 > 0$ and $\alpha_0^2 < 0$, then utility rankings with $U_1(1,1) - U_1(1,0) > 0$ and $U_2(0,1) - U_2(0,0) < 0$ cannot occur. Representation (1.2) is similar to that in the "simultaneous-move" games as proposed by Bresnahan and Reiss (1991). Let be:

$$\beta_p = \beta_1^p - \beta_0^p, \quad \alpha_p = \alpha_1^p - \alpha_0^p, \quad \varepsilon_p = \varepsilon_1^p - \varepsilon_0^p$$

for $p = \{1, 2\}$. Under this structure of the game we can define an equilibrium concept.

a) Nash Equilibrium

Following Bjorn and Vuong (1984), player is assumed to maximize their utility function, given the action of the other player; this means that players adjust their actions until decisions are mutually consistent. Let $\{(l,k) = (0,1)\}$, then, (k,l) is a Nash equilibrium if:

$$U_1(k,l) > U_1(1-k,l) \quad \text{and} \quad U_2(k,l) > U_2(k,1-l)$$

Then, NE will depend on signs of the following utility differences: *reaction functions*.

$$\begin{aligned}
&\textbf{For Player 1}\\
U_1(1,1) - U_1(0,1) &= X'\beta_1 + \alpha_1 + \varepsilon_1\\
U_1(1,0) - U_1(0,0) &= X'\beta_1 + \varepsilon_1\\
&\textbf{For Player 2}\\
U_2(1,1) - U_2(0,1) &= X'\beta_2 + \alpha_2 + \varepsilon_2\\
U_2(1,0) - U_2(0,0) &= X'\beta_2 + \varepsilon_2 \qquad (1.4)
\end{aligned}$$

There are 16 different combinations which will determine the equilibrium. In some cases there are two NE (multiple equilibria); whereas in others don't exist (remember the blank square in Figure 2 named "Mixed"). In order to identify the parameters of the model, when there are multiple equilibria or none, Bjorn and Vuong (1984) assumed that players choose one of the equilibria at random, such that each equilibrium is chosen with equal probabilities. In the other case in which there is not action chosen (blank square in Figure 2), players are assumed to choose one of the four allocation with equal probabilities. This has been criticized as *ad hoc* assumptions.

In terms of identification, notice that α_1, α_2, β_1, β_2 are identified, but α_0^1, α_1^1, α_0^2, α_1^2, and β_0^1, β_1^1, β_0^2, β_1^2 are not identified separately.

Finally, we want to know the likelihood contributions of each allocation:

$$\{(1,1),(1,0),(0,1),(0,0)\}$$

Under certain regularity conditions[6] and assumptions made above, we have that for $(0,1)\times(0,1)$, there are only 4 possible *reaction functions* for each player:

From (1.3), define:

For Player 1

$$U_1^1 = \begin{cases} U_1(1,1) - U_1(1,0) > 0 & \text{and} \\ U_1(0,1) - U_1(0,0) > 0 & \Leftrightarrow \varepsilon_1 > -X'\beta_1 - min(0,\alpha_1) \end{cases}$$
$$U_1^2 = \begin{cases} U_1(1,1) - U_1(1,0) > 0 & \text{and} \\ U_1(0,1) - U_1(0,0) < 0 & \Leftrightarrow -X'\beta_1 - \alpha_1 < \varepsilon_1 < -X'\beta_1 \end{cases}$$
$$U_1^3 = \begin{cases} U_1(1,1) - U_1(1,0) < 0 & \text{and} \\ U_1(0,1) - U_1(0,0) > 0 & \Leftrightarrow -X'\beta_1 < \varepsilon_1 < -X'\beta_1 - \alpha_1 \end{cases}$$
$$U_1^4 = \begin{cases} U_1(1,1) - U_1(1,0) < 0 & \text{and} \\ U_1(0,1) - U_1(0,0) < 0 & \Leftrightarrow \varepsilon_1 < -X'\beta_1 - max(0,\alpha_1). \end{cases}$$

For Player 2

$$U_2^1 = \begin{cases} U_2(1,1) - U_2(1,0) > 0 & \text{and} \\ U_2(0,1) - U_2(0,0) > 0 & \Leftrightarrow \varepsilon_2 > -X'\beta_2 - min(0,\alpha_2) \end{cases}$$
$$U_2^2 = \begin{cases} U_2(1,1) - U_2(1,0) > 0 & \text{and} \\ U_2(0,1) - U_2(0,0) < 0 & \Leftrightarrow -X'\beta_2 - \alpha_2 < \varepsilon_2 < -X'\beta_2 \end{cases}$$
$$U_2^3 = \begin{cases} U_2(1,1) - U_2(1,0) < 0 & \text{and} \\ U_2(0,1) - U_2(0,0) > 0 & \Leftrightarrow -X'\beta_2 < \varepsilon_2 < -X'\beta_2 - \alpha_2 \end{cases}$$
$$U_2^4 = \begin{cases} U_2(1,1) - U_2(1,0) < 0 & \text{and} \\ U_2(0,1) - U_2(0,0) < 0 & \Leftrightarrow \varepsilon_2 < -X'\beta_2 - max(0,\alpha_2). \end{cases}$$

[6]See Bjorn and Vuong (1984).

Nash Model

Players	U_1^1	U_1^2	U_1^3	U_1^4
U_2^1	(1,1)	(1,1)	(0,1)	(0,1)
U_2^2	(1,1)	(1,1) or (0,0)	No NE	(0,0)
U_2^3	(1,0)	No NE	(1,0) or (0,1)	(0,1)
U_2^4	(1,0)	(0,0)	(1,0)	(0,0)

Notice that (U_2^2, U_1^2) and (U_2^3, U_1^3) represent a *multiple equilibria*; whereas, (U_2^2, U_1^3) and (U_2^3, U_1^2), represent no equilibria. Likelihood contributions can be derived straightforwardly (see Bjorn and Vuong, 1984). Estimation of $\boldsymbol{\theta} = \{\alpha_1, \alpha_2, \beta_1, \beta_2\}$ can be made using likelihood framework.

b) Stackelberg Equilibrium

Following Bjorn and Vuong (1985), in a Stackelberg game the role of the players is asymmetric. One of the players, *the leader*, is assumed to maximize his utility anticipating the reaction of the other player, *the follower*. Formally, allocation (l, k) is SE with Player 1 being the *leader* and Player 2, the *follower* if

$$\begin{aligned} &U_2(k,l) > U_2(k,1-l) \qquad \text{and} \quad U_1(k,l) > U_1(1-k,l) \\ &U_2(1-k,l) > U_2(1-k,1-l) \end{aligned}$$

or

$$\begin{aligned} &U_2(k,l) > U_2(k,1-l) \qquad \text{and} \quad U_1(k,l) > U_1(1-k,1-l) \\ &U_2(1-k,l) < U_2(1-k,1-l) \end{aligned}$$

Stackelberg Model

Player 2	Player 1		
U_2^1	$U_1(1,1) - U_1(1,0) > 0 \Leftrightarrow$	$\varepsilon_1 > -X'\beta_1 - \alpha_1$	(1,1) is SE
U_2^1	$U_1(1,1) - U_1(1,0) < 0 \Leftrightarrow$	$\varepsilon_1 < -X'\beta_1 - \alpha_1$	(0,1) is SE
U_2^2	$U_1(1,1) - U_1(0,0) > 0 \Leftrightarrow$	$\varepsilon_1 > -X'\beta_1 - \alpha_1^1$	(1,1) is SE
U_2^2	$U_1(1,1) - U_1(0,0) < 0 \Leftrightarrow$	$\varepsilon_1 < -X'\beta_1 - \alpha_1^1$	(0,0) is SE
U_2^3	$U_1(1,0) - U_1(0,1) > 0 \Leftrightarrow$	$\varepsilon_1 > -X'\beta_1 + \alpha_0^1$	(1,1) is SE
U_2^3	$U_1(1,0) - U_1(0,1) < 0 \Leftrightarrow$	$\varepsilon_1 < -X'\beta_1 + \alpha_0^1$	(0,1) is SE
U_2^4	$U_1(1,0) - U_1(0,0) > 0 \Leftrightarrow$	$\varepsilon_1 > -X'\beta_1$	(1,0) is SE
U_2^4	$U_1(1,0) - U_1(0,0) < 0 \Leftrightarrow$	$\varepsilon_1 < -X'\beta_1$	(0,0) is SE

Where, U_2^i, $i = \{1,2,3,4\}$, were defined above. As opposite of NE, SE is always defined uniquely. Notice that β_1, α_1^1, and α_0^1 are identified by the *leader*; whereas β_2 and α_2 are only identified by the *follower*. Likelihood function can be derived straightforward from the last table and a certain distributional assumptions over $(\varepsilon_1, \varepsilon_2)$, see Bjorn and Vuong, 1985.

c) Pareto Optimality

Following Koreman (1994), Pareto optimality can be reach if

$$\Big(U_1(k,l) > U_1(k,1-l) \quad \text{or} \quad U_1(k,l) > U_2(k,1-l)\Big)$$
$$\text{and}$$
$$\Big(U_1(k,l) > U_1(k,1-l) \quad \text{or} \quad U_1(k,l) > U_2(k,1-l)\Big)$$
$$\text{and}$$
$$\Big(U_1(k,l) > U_1(k,1-l) \quad \text{or} \quad U_1(k,l) > U_2(k,1-l)\Big)$$

As before, we assume that the players choose one of the Pareto optimal allocations at random, such that each Pareto optimal allocation is chosen with equal probabilities. Similar to NE and SE, can be derived the likelihood contributions, see Koreman, 1994.

d) Nonlinear Simultaneous Equations Model

Following Tamer (2003), we can say that his paper contributes to the literature on inference in nonlinear simultaneous equations model. From model describe by

$$\begin{aligned} y_1^* &= X_1'\beta_1 + y_2\alpha_1 + \varepsilon_1 \\ y_2^* &= X_2'\beta_2 + y_1\alpha_2 + \varepsilon_2 \end{aligned}$$

$$y_p = \begin{cases} 1, & \text{if } y_p^* > 0; \\ 0, & \text{otherwise.} \end{cases}$$

For $p = \{1,2\}$. If the econometrist knows signs of $\alpha's$, as seen, there are two situations:

d.1.) $\alpha_1 < 0$, $\alpha_2 < 0$

If ε_p has enough support, the theoretical game admits *multiple equilibria*. In this case for $-X_p'\beta_p \leq \varepsilon_p \leq -X_p'\beta_p - \alpha_p$, $(p = 1,2)$, the econometric model predicts either (0,1) or (1,0). then, the model provides the following inequality restrictions on conditional regressions: For,

$$\boldsymbol{\theta} = (\beta_1, \beta_2, \alpha_1, \alpha_2, \Omega)$$

$$\begin{aligned} P_1(\mathbf{X}, \boldsymbol{\theta}) &= Pr[(0,0)|\mathbf{x}] = Pr(\varepsilon_1 < -X_1'\beta_1; \varepsilon_2 < -X_2'\beta_2) \\ P_2(\mathbf{X}, \boldsymbol{\theta}) &= Pr[(1,1)|\mathbf{x}] = Pr(\varepsilon_1 \geq -X_1'\beta_1 - \alpha_1; \varepsilon_1 \geq -X_2'\beta_2 - \alpha_2) \\ P_3(\mathbf{X}, \boldsymbol{\theta}) &\leq Pr[(0,1)|\mathbf{x}] \leq P_4(\mathbf{X}, \boldsymbol{\theta}) \end{aligned} \tag{1.5}$$

where

$$\begin{aligned} P_3(\mathbf{X}, \boldsymbol{\theta}) &= Pr(\varepsilon_1 < -X_1'\beta_1 - \alpha_1; \varepsilon_2 > -X_2'\beta_2 - \alpha_2) + \\ &+ Pr(\varepsilon_1 < -X_1'\beta_1; -X_2'\beta_2 < \varepsilon_2 < -X_2'\beta_2 - \alpha_2) \\ P_4(\mathbf{X}, \boldsymbol{\theta}) &= Pr(\varepsilon_1 < -X_1'\beta_1 - \alpha_1; \varepsilon_2 > -X_2'\beta_2) \end{aligned}$$

The upper an lower probabilities on the (1,0) outcome are similar. The bound on the conditional probabilities provided by the incomplete model are usually much tighter than the ones obtained from models that treat the (0,1) and (1,0) outcomes as one event. The bound provided by these models on the (0,1) outcome is

$$0 \leq Pr[(0,1)|\mathbf{x}] \leq 1 - P_1(\mathbf{X},\boldsymbol{\theta}) - P_2(\mathbf{X},\boldsymbol{\theta})$$

If we allow *mixed strategies* in the underlying game, then the restrictions provided by the model will be

$$\begin{aligned} P_1(\mathbf{X},\boldsymbol{\theta}) &\leq Pr[(0,0)|\mathbf{x}] \leq P_1(\mathbf{X},\boldsymbol{\theta}) + P_{square}(\mathbf{X},\boldsymbol{\theta}) \\ P_2(\mathbf{X},\boldsymbol{\theta}) &\leq Pr[(1,1)|\mathbf{x}] \leq P_2(\mathbf{X},\boldsymbol{\theta}) + P_{square}(\mathbf{X},\boldsymbol{\theta}) \\ P_3(\mathbf{X},\boldsymbol{\theta}) &\leq Pr[(0,1)|\mathbf{x}] \leq P_4(\mathbf{X},\boldsymbol{\theta}) \end{aligned} \tag{1.6}$$

where

$$P_{square}(\mathbf{X},\boldsymbol{\theta}) = Pr[-X_1'\beta_1 < \varepsilon_1 < -X_1'\beta_1 - \alpha_1; -X_2'\beta_2 < \varepsilon_2 < -X_2'\beta_2 - \alpha_2|\mathbf{x}]$$

Square can be seen in Figure 3.

d.1.) $\alpha_1 > 0$, $\alpha_2 < 0$

Blanket square in Figure 2, represents that some values of the exogenous variables, either of the four outcomes is likely. In this square, there is no equilibrium in pure strategies: each player is indifferent between choosing 1 or 0 given that the other is randomizing. This maps into the model having the following restrictions:

$$\begin{aligned} P_1(\mathbf{X},\boldsymbol{\theta}) &\leq Pr[(0,0)|\mathbf{x}] \leq P_1(\mathbf{X},\boldsymbol{\theta}) + P_{square}(\mathbf{X},\boldsymbol{\theta}) \\ P_2(\mathbf{X},\boldsymbol{\theta}) &\leq Pr[(1,1)|\mathbf{x}] \leq P_2(\mathbf{X},\boldsymbol{\theta}) + P_{square}(\mathbf{X},\boldsymbol{\theta}) \\ P_3'(\mathbf{X},\boldsymbol{\theta}) &\leq Pr[(0,1)|\mathbf{x}] \leq P_3'(\mathbf{X},\boldsymbol{\theta}) + P_{square}(\mathbf{X},\boldsymbol{\theta}) \end{aligned} \tag{1.7}$$

Where P_1 and P_2 are the same as above and

$$P_3'(\mathbf{X}, \boldsymbol{\theta}) = Pr[\varepsilon_1 \geq -X_1'\beta_1; \varepsilon_2 < -X_2'\beta_2 - \alpha_2]$$

and

$$P_{square}(\mathbf{X}, \boldsymbol{\theta}) = Pr[-X_1'\beta_1 - \alpha_1 \leq \varepsilon_1 \leq -X_1'\beta_1; -X_2'\beta_2 \leq \varepsilon_2 \leq -X_2'\beta_2 - \alpha_2|\mathbf{x}]$$

Maximum Likelihood Estimator ($\alpha_1 < 0$ and $\alpha_2 < 0$): Multiple equilibrium is allowed. This maximum likelihood estimator can be used to consistently estimate parameters $\boldsymbol{\theta}$. It is considered the model in which there are only three outcomes: (0,0), (1,1) and (1,0) and (0,1). Given the outcome probabilities given in (1.5). Assumptions A1-A3 hold, and some regularity conditions[7], we have

$$\begin{aligned} \mathcal{L}_{ML}(b) &= \frac{1}{N}\sum_{i=1}^{N}\Big[y_{i1}y_{i2}\log(P_1(\mathbf{X}, \mathbf{b})) + (1-y_{i1})(1-y_{i2})\log(P_2(\mathbf{X}, \mathbf{b})) \\ &+ ((1-y_{i1})y_{i2} + y_{i1}(1-y_{i2}))\log(1 - P_1(\mathbf{X}, \mathbf{b}) - P_2(\mathbf{X}, \mathbf{b}))\Big] \end{aligned} \tag{1.8}$$

where P_1 and P_2 are defined in (1.5). Using Maximum Likelihood Estimation techniques, the covariance matrix of the above-modified likelihood is

$$\Omega_{ML} = E\left[\frac{\partial P_1 \partial P_1'}{P_1} + \frac{\partial P_2 \partial P_2'}{P_2} + \frac{(\partial P_1 + \partial P_2)(\partial P_1 + \partial P_2)'}{1 - P_1 - P_2}\right]^{-1}$$

where ∂P_1 and ∂P_2 are the derivative vectors of the functions $P_1(\mathbf{X}, \mathbf{b})$, and $P_2(\mathbf{X}, \mathbf{b})$ respectively, with respect to $\mathbf{b}$ evaluated in the true parameters $\boldsymbol{\theta}_0$.

Incomplete Information

Simultaneous 2×2 game in its normal representation becomes.

[7]See Tamer, 2003.

Figure 5

A simple 2 x 2 game

		PLAYER 2 Y=1	PLAYER 2 Y=0
PLAYER 1	Y=1	$t_1 + \alpha_1 \pi_2, t_2 + \alpha_2 \pi_1$	$t_1, 0$
	Y=0	$0, t_2$	0,0

Following Aradillas-López (2008), we will incorporate the assumption of *incomplete information*. This implies that players don't know payoff function of the other one, they should guess about it in order to take their decisions. In this context, we assume that players do not know Y_{-p}, instead of that, they incorporate beliefs in their payoff function as follows

$$\pi_{-p} = Pr_p(Y_{-p} = 1 | Y_p = 1, X)$$

Which is the player p belief.

Equilibrium beliefs and actions

Given the normal-form of the game (Figure 5) and assumptions (A1)-(A2) (just modified by its *incomplete informational* assumption, i.e., players should guess other players actions), players' optimal actions are given by

$$\begin{aligned} Y_1 &= 1\{X_1'\beta_1 + \alpha_1 \underbrace{Pr_1(Y_2 = 1 | Y_1 = 1, \mathbf{X})}_{\pi_2 = \text{Player 1's beliefs}} - \varepsilon_1 \geq 0\} \\ Y_2 &= 1\{X_2'\beta_2 + \alpha_2 \underbrace{Pr_2(Y_1 = 1 | Y_2 = 1, \mathbf{X})}_{\pi_1 = \text{Player 2's beliefs}} - \varepsilon_2 \geq 0\} \end{aligned} \tag{1.9}$$

Where $1\{A\}$ is the indicator function: $1\{A\} = 1$ if A is true, 0 otherwise.

Then, husband and wife will use pure-strategy based on (1.9). Our goal is to find a pair of *self-consistence* equilibrium believes that satisfy (1.9). Let be a pair of scalars $\pi_1, \pi_2 \in \mathbb{R}^2$ and define:

$$\begin{aligned}
\varphi_1(\pi_1, \pi_2; \boldsymbol{\theta}, \mathbf{X}) &= \pi_1 - \frac{G_{1,2}(X_1'\beta_1 + \alpha_1\pi_2, X_2'\beta_2 + \alpha_2\pi_1; \rho)}{G_2(X_2'\beta_2 + \alpha_2\pi_1)} \\
\varphi_2(\pi_1, \pi_2; \boldsymbol{\theta}, \mathbf{X}) &= \pi_2 - \frac{G_{1,2}(X_1'\beta_1 + \alpha_1\pi_2, X_2'\beta_2 + \alpha_2\pi_1; \rho)}{G_1(X_1'\beta_1 + \alpha_1\pi_2)} \qquad (1.10) \\
\varphi(\pi_1, \pi_2; \boldsymbol{\theta}, \mathbf{X}) &= (\varphi_1(\pi_1, \pi_2; \boldsymbol{\theta}, \mathbf{X}), \varphi_2(\pi_1, \pi_2; \boldsymbol{\theta}, \mathbf{X}))
\end{aligned}$$

Proposition 1

This is a version of Proposition 1 in Aradillas-Lopez (2008). Given a fix realization of $\boldsymbol{x} \in \mathbb{S}(\mathbf{X})$, ρ and the value of $\boldsymbol{\theta}$, then, $\varphi(\pi_1, \pi_2; \boldsymbol{\theta}, \mathbf{X})$ can be considered as a function of $(\pi_1\pi_2)$. If assumptions (A1) and (A2) are satisfied, players' beliefs are deterministic given $\boldsymbol{x}$ and a pair of *self-consistent* beliefs that satisfy (1.9) must solve for (π_1, π_2) the system

$$\varphi(\pi_1, \pi_2; \boldsymbol{\theta}, \mathbf{X}) = \mathbf{0} \qquad (1.11)$$

For a given values of $\mathbf{x}$, ρ and $\boldsymbol{\theta}$, roughly speaking, we will say that "there exists an equilibrium"[8] if there exists a pair of (π_1, π_2) such that (1.11) is satisfied. A detail analysis of existence, cardinality and uniqueness of the equilibria, can be found in Aradillas-Lopez (2008).

Estimation of π_1 and π_2

If there exist a *unique* solution of the equilibrium system (1.11) for a given values of $\boldsymbol{x}$, ρ and $\boldsymbol{\theta}$, it will be denoted as $\hat{\pi}(\boldsymbol{\theta}, \boldsymbol{x}) \equiv \big(\hat{\pi}_1(\boldsymbol{\theta}, \boldsymbol{x}), \hat{\pi}_2(\boldsymbol{\theta}, \boldsymbol{x})\big)$.

[8] Here is allow multiple equilibria. If that were the case, we could choose the smalest $\hat{\pi}_1$, and uses it for the analysis. (see Aradillas-Lopez, 2008).

We are interested in to find $\pi^*(\boldsymbol{\theta}, \boldsymbol{x})$ (self-consistent beliefs). We estimate it as follows: take a pair of $(\pi_1, \pi_2) \in \mathrm{R}^2$. Let be

$$Q(\pi; \boldsymbol{\theta}, \boldsymbol{x}) = -\varphi(\pi; \boldsymbol{\theta}, \boldsymbol{x})' \varphi(\pi; \boldsymbol{\theta}, \boldsymbol{x}) \tag{1.12}$$

With $\varphi(\pi; \boldsymbol{\theta}, \boldsymbol{x})$, defined in (1.10), so we maximize

$$\max_{\pi \in [0,1]^2} Q(\pi; \boldsymbol{\theta}, \boldsymbol{x}) \tag{1.13}$$

Which give us $\hat{\pi}(\boldsymbol{\theta}, \boldsymbol{x}) = \big(\hat{\pi}_1(\boldsymbol{\theta}, \boldsymbol{x}), \hat{\pi}_2(\boldsymbol{\theta}, \boldsymbol{x})\big)$, estimation of players' *self-consistent* beliefs.

Estimation of $\boldsymbol{\theta}$

Introducing some notation: let be $t_1 = X_1'\beta_1 + \alpha_1\pi_2$ and $t_2 = X_2'\beta_2 + \alpha_2\pi_1$, where $(\pi_1, \pi_2) \in \mathbb{R}^2$. Let define,

$$\begin{aligned}
P_{1,1}(\mathbf{X}, \boldsymbol{\theta}, \pi) &= \Pr(t_1 \geq \varepsilon_1, t_2 \geq \varepsilon_2 | \mathbf{X}; \rho) \\
P_{1,0}(\mathbf{X}, \boldsymbol{\theta}, \pi) &= \Pr(t_1 \geq \varepsilon_1, t_2 < \varepsilon_2 | \mathbf{X}; \rho) \\
P_{0,1}(\mathbf{X}, \boldsymbol{\theta}, \pi) &= \Pr(t_1 < \varepsilon_1, t_2 \geq \varepsilon_2 | \mathbf{X}; \rho) \\
P_{0,0}(\mathbf{X}, \boldsymbol{\theta}, \pi) &= \Pr(t_1 < \varepsilon_1, t_2 < \varepsilon_2 | \mathbf{X}; \rho)
\end{aligned} \tag{1.14}$$

Define $\mathbf{W}$=($\mathbf{Y}$,$\mathbf{X}$), and

$$\begin{aligned}
\mathscr{L}_{1,1}(\mathbf{X},\boldsymbol{\theta},\pi) &\equiv \log P_{1,1}(\mathbf{X},\boldsymbol{\theta},\pi) \\
\mathscr{L}_{1,0}(\mathbf{X},\boldsymbol{\theta},\pi) &\equiv \log P_{1,0}(\mathbf{X},\boldsymbol{\theta},\pi) \\
\mathscr{L}_{0,1}(\mathbf{X},\boldsymbol{\theta},\pi) &\equiv \log P_{0,1}(\mathbf{X},\boldsymbol{\theta},\pi) \qquad (1.15) \\
\mathscr{L}_{0,0}(\mathbf{X},\boldsymbol{\theta},\pi) &\equiv \log P_{0,0}(\mathbf{X},\boldsymbol{\theta},\pi)
\end{aligned}$$

Then, the conditional log-likelihood of $\mathbf{Y}$ given $\mathbf{X}$ is given by:

$$\begin{aligned}
\mathscr{L}(\mathbf{W},\boldsymbol{\theta},\pi) &= Y_1Y_2\mathscr{L}_{1,1}(\mathbf{X},\boldsymbol{\theta},\pi) + (1-Y_1)Y_2\mathscr{L}_{1,0}(\mathbf{X},\boldsymbol{\theta},\pi) + \\
&\quad +Y_1(1-Y_2)\mathscr{L}_{1,0}(\mathbf{X},\boldsymbol{\theta},\pi) + \qquad (1.16) \\
&\quad +(1-Y_1)(1-Y_2)\mathscr{L}_{0,0}(\mathbf{X},\boldsymbol{\theta},\pi)
\end{aligned}$$

Using the self-consistent equilibrium believes which were found using (1.11), $\hat{\pi}(\mathbf{X},\boldsymbol{\theta})$, then we find $\boldsymbol{\theta}$ by solving:

$$\max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \frac{1}{N}\sum_{i=1}^{N}\mathscr{L}(W_i,\boldsymbol{\theta},\hat{\pi}(\mathbf{X},\boldsymbol{\theta})) \qquad (1.17)$$

Where $i = \{1, 2, ..., N\}$ games.

Matrix of variances and covariances can be found using maximum likelihood estimation techniques (see Amemiya, 1985). Then

$$\Omega_{MLE} = -\left[E\frac{\partial^2\mathscr{L}(\mathbf{W},\boldsymbol{\theta},\pi)}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}\right]^{-1} \qquad (1.18)$$

Which is known as Cramer-Rao lower bound.

1.2.2 Dynamic Games

Empirical industrial organization economists have proposed estimates for dynamic games of incomplete information. In these models, agents choose from a finite number actions and maximize expected discounted utility in a Markov perfect equilibrium. Some econometric methods have estimated the probability distribution of agents actions in a first stage. In a second step, a finite vector of parameters of the period, return function are estimated. Agents are forward looking and maximize expected discounted utility. Like the models surveyed in Rust (1994) or studied in Keane and Wolpin (1997).

Nonetheless, there is another way to model dynamic games: agents interact strategically and play a Markov perfect equilibrium to a dynamic game. See Pakes, Ostrovsky and Berry (2004), Aguirregabiria and Mira (2002), Pesendorfer and Schmidt-Dengler (2003) and Bajari, Benkard and Levin (2003). Substantive applications of dynamic games estimators include Jenkins, Liu, McFadden, and Matzkin (2004) to the browser war.

Bajari and Hong (2006) have developed a semiparametric estimators for dynamic games allowing for continuous state varaibles and a nonparametric first stage, similar to models discussed by Pakes, Ostrovsky and Berry (2004), Aguirregabiria and Mira (2002), Pesendorfer and Schmidt-Dengler (2003) and Bajari, Benkard and Levin (2003).

Nonparametric identification results for dynamic discrete games are developed by Aguirregabiria and Mira (2002) and Pesendorfer and Schmidt-Dengler (2003) in the context of discrete state space models. Relatively recent works by Heckman and Navarro (2005) and Aguirregabiria (2005) present identification results for dynamic discrete choice models allowing for continuous state variables.

1.3 Labor Participation: Wife and Husband Decision Game

Let be player 1, *husband*; and player 2 *wife*, in a married couple which play a game with *complete information* in order to decide their participation or not in the labor market. Using the game structure given in Figure 1, Y=1 means *participate* in the labor market, and Y=0, *don't participate.* As usual, upper-case will denote random variables and lower-case, particular realizations of this random variables. As mentioned, $\mathbb{S}(U)$, represents the support of the random variable U; and the subscript $p \in \{1, 2\}$ to denote a particular player, and $-p$ to denote the opponent. Strategic parameters will be represented as $(\alpha_1, \alpha_2) \in \mathbb{R}^2$. This parameters summarize the interaction effect between the players' actions.

Following Bjorn and Vuong (1984)[9] we assume that α_1 (husband's interaction coefficient) is negative. This means that husband would be affected if his wife decides to work (because of social considerations, among others). At the same time, it is assumed that α_2 is negative too. The fact that both (α_1, α_2) were negative, i.e., the assumption of is hurt for both participate in the labor market is reasonable because they would be affected because they could have a child, or they would be apart when work, etc.; and means that *multiple equilibria* is allowed in the context of *complete information* environment.

As mentioned, this kind of models have been widely studied in econometrics, Bresnahan and Reis (1991) and by Heckman (1978) in his seminal paper (models with structural shifts parameters). Blundell and Smith (1993, 1994) on female labour supply and Schmidt (1981) on simultaneity in bivariate econometrics models.

When multiple equilibrium is allowed, some authors have made *ad-hoc* assumptions: decision process is assumed that come from a "single equilibrium concept". As

[9]This is the assumptions which indicates that econometrist knows the signs of $\alpha's$. Bjorn and Vuong (1984) found that interaction parameters in the U.S. where negative.

it was mentioned, pioneer work is Bjorn and Vuong (1984, 1985). Kooreman (1994), estimated and compared some microeconometric models for simultaneous discrete endogenous variables; he used data on the joint labor force participation decisions of husbands and wives in a sample of Dutch households, under the assumption that outcomes came from a Nash Equilibrium, Stackelberg Equilibrium and Pareto optimallity. Then, they make simplifying assumptions to respond to the nonuniqueness problem without invoking the coherency condition.

On the other hand, Tamer (2003), proposed a parametric and nonparametric estimator without invoking this coherency condition nor imposing *ad-hoc* assumptions to avoid multiplicity equilibria. Our model go forward because not only allows multiplicity equilibrium, but also designs an equilibrium selection mechanism.

1.3.1 Equilibrium Selection Mechanism

As $\alpha_1 \leq 0$, and $\alpha_2 \leq 0$, we have multiple equilibria. Then, players should decide how to choose the optimal equilibrium. Equilibria will be ordered according to the husband's probability to work. In the square area in Figure 3, there are three equilibriums: (0,1), mixed, and (1,0). In the first one, husband will work with probability zero; in the second one he will work with probability $\Pi_1 = -\frac{t_2}{\alpha_2} \in [0,1]$; in the third one, he will work with probability one. Then, it will be considered an *ordered response approach* using a linear index.

Let be,

$$\mathbf{W}'\gamma + \eta \tag{1.19}$$

the linear index where $\mathbf{W}$ are observable characteristics using *exclusion restriction*, which means that we should include some characteristics that are not included in $\mathbf{X}_1$ or $\mathbf{X}_2$. $\gamma \in \mathrm{R}^{k_2}$ is a vector of parameters and η is unobservable vector. $k_2 \in \mathrm{Z}$.

Figure 6

Linear Index

where μ_a and μ_b are *threshold parameters*, such that $\mu_b \geq \mu_a$. If $\gamma_i > 0$ contributes husband to work. Finally, η is assumed independent of $(\varepsilon_1, \varepsilon_2)$[10].

According to Figure 3, regions have the following specific expressions:

$\mathbf{R}_{(sqr)}$: $0 \leq t_2 \leq -\alpha_2$ and $0 \leq t1 \leq -\alpha_1$

$\mathbf{R}_{(1,1)}$: $t_2 > -\alpha_2$ and $t1 > -\alpha_1$

$\mathbf{R}_{(0,0)}$: $t_2 < 0$ and $t_1 < 0$

$\mathbf{R}_{(0,1)}$: $t_2 > 0$ and $t_1 < 0$

$\bigcup$

$t_2 > -\alpha_2$ and $0 \leq t_1 \leq -\alpha_1$

$\mathbf{R}_{(1,0)}$: $t_2 < 0$ and $t_1 > 0$

$\bigcup$

$0 \leq t_2 \leq -\alpha_2$ and $t_1 > -\alpha_1$

Then, we need to construct specific expressions for Pr(1,1), Pr(0,0), Pr(0,1) and Pr(1,0). But first I will describe informational assumptions.

1.4 Estimation

Assumption A4 (**Researcher**) Main assumption made here is that distributions of $(\varepsilon_1, \varepsilon_2) \sim G_{1,2}(.,.;\rho)$, and $\eta \sim F(\bullet)$ are assumed to be known.

Here, it will be used the Farlie-Gumbel-Morgesten families of joint distributions (see Johnson, et. al. 1999), then $G_{1,2}(.,.;\rho)$ can be expressed as follows:

[10] This could be relaxed.

$$G_{1,2}(\epsilon_1, \epsilon_2; \rho) = G_1(\epsilon_1)G_2(\epsilon_2) \times \left[1 + \rho(1 - G_1(\epsilon_1))(1 - G_2(\epsilon_2))\right] \tag{1.20}$$

where

$$G(\epsilon_p) = \frac{e^{\epsilon_p}}{1 + e^{\epsilon_p}} \tag{1.21}$$

$G(\bullet)$ is the logistic cdf.

At the same time, we assume that η has a logistic cdf.

$$F(\eta) = \frac{e^{\eta}}{1 + e^{\eta}} \tag{1.22}$$

Assumption A5 (**Researcher**) The econometrician has in hand an iid sample of N games described by the assumptions (A1)-(A4). He observes $(Y_n, X_n, W_n)_{n=1}^{N}$, and uses the joint cdf $G_{1,2}(.,.;\rho)$ described above with $\rho \leq |1|$, and the logistical, $F(\bullet)$, for η, in order to identify all parameters.

1.4.1 Probabilities

Under the assumptions (A1)-(A5), probability functions for each pair of actions $\{(1,1),(0,0),(1,0),(0,1)\}$ are defined as follow.

Probability of (1,0)

Let be $1\{(1,0)\}$ the indicator function for the simultaneous actions: $Y_1 = 1$ and $Y_2 = 0$. From the Figure 3, and including the *selection mechanism*, we have:

$$\begin{aligned} 1\{(1,0)\} = \quad & 1\{(t_1,t_2) \in \mathbf{R}_{(1,0)}\} + \\ & 1\{(t_1,t_2) \in \mathbf{R}_{(square)} \text{ and } W'\gamma + \eta > \mu_b\} + \\ & 1\{(t_1,t_2) \in \mathbf{R}_{(square)} \text{ and } \mu_a \leq W'\gamma + \eta \leq \mu_b\} \\ & 1\{U_2 \leq \Pi_2\}1\{U_1 > \Pi_1\} \end{aligned} \tag{1.23}$$

where U_1 and U_2 are uniform random variables in [0,1], independent from all random variables in the game and between them. $1\{U_2 \leq \Pi_2\}$, and $1\{U_1 > \Pi_1\}$ are called *randomization device*. Notice that all possibilities have been considered, because regions are mutually exclusive from each other.

Now, the conditional probability of $\{(1,0)\}$ given X_1, X_2, W, ε_1, ε_2, and η is:

$$\begin{aligned} Pr[\{(1,0)\}|X_1, X_2, W, \varepsilon_1, \varepsilon_2, \eta] = \quad & 1\{(t_1, t_2) \in \mathbf{R}_{(1,0)}\} + \\ & 1\{(t_1, t_2) \in \mathbf{R}_{(square)} \text{ and } W'\gamma + \eta > \mu_b\} + \\ & 1\{(t_1, t_2) \in \mathbf{R}_{(square)} \text{ and } \mu_a \leq W'\gamma + \eta \leq \mu_b\} \\ & \Pi_2(1 - \Pi_1) \end{aligned} \tag{1.24}$$

Finally, integrating over $\varepsilon_1, \varepsilon_2$,and η, we have:

$$\begin{aligned} \mathscr{P}_{(1,0)} = Pr(\{1,0\}|X_1, X_2, W) = & \\ \int_{\varepsilon_1}\int_{\varepsilon_2}\int_{\eta} \quad & \Big[1\{(t_1, t_2) \in \mathbf{R}_{(1,0)}\} + \\ & 1\{(t_1, t_2) \in \mathbf{R}_{(square)} \text{ and } W'\gamma + \eta > \mu_b\} + \\ & 1\{(t_1, t_2) \in \mathbf{R}_{(square)} \text{ and } \mu_a \leq W'\gamma + \eta \leq \mu_b\} \\ & \Pi_2(1 - \Pi_1)\Big] \; g_{1,2}(\epsilon_1, \epsilon_2; \rho) \; f(\eta) \; d\eta \; d\varepsilon_2 \; d\varepsilon_1 \end{aligned} \tag{1.25}$$

Parallel structure was used in order to find {(1,1)}, {(0,0)}, and {(0,1)} probabilities, so we get:

Probability of (1,1)

$$\begin{aligned} \mathscr{P}_{(1,1)} = Pr(\{1,1\}|X_1, X_2, W) = & \\ \int_{\varepsilon_1}\int_{\varepsilon_2}\int_{\eta} \quad & \Big[1\{(t_1, t_2) \in \mathbf{R}_{(1,1)}\} + \\ & 1\{(t_1, t_2) \in \mathbf{R}_{(square)} \text{ and } \mu_a \leq W'\gamma + \eta \leq \mu_b\} \\ & \pi_2\pi_1\Big] \; g_{1,2}(\epsilon_1, \epsilon_2; \rho) \; f(\eta) \; d\eta \; d\varepsilon_2 \; d\varepsilon_1 \end{aligned} \tag{1.26}$$

Probability of (0,0)

$$\mathscr{P}_{(0,0)} = Pr(\{0,0\}|X_1, X_2, W) =$$

$$\int_{\varepsilon_1}\int_{\varepsilon_2}\int_{\eta} \Big[1\{(t_1,t_2) \in \mathbf{R}_{(0,0)}\} + \\ 1\{(t_1,t_2) \in \mathbf{R}_{(square)} \text{ and } \mu_a \leq W'\gamma + \eta \leq \mu_b\} \\ (1-\pi_2)(1-\pi_1)\Big]\ g_{1,2}(\epsilon_1, \epsilon_2; \rho)\ f(\eta)\ d\eta\ d\varepsilon_2\ d\varepsilon_1 \tag{1.27}$$

Probability of (0,1)

$$\mathscr{P}_{(0,1)} = Pr(\{0,1\}|X_1, X_2, W) =$$

$$\int_{\varepsilon_1}\int_{\varepsilon_2}\int_{\eta} \Big[1\{(t_1,t_2) \in \mathbf{R}_{(0,1)}\} + \\ 1\{(t_1,t_2) \in \mathbf{R}_{(square)} \text{ and } \mu_a \leq W'\gamma + \eta\} + \\ 1\{(t_1,t_2) \in \mathbf{R}_{(square)} \text{ and } \mu_a \leq W'\gamma + \eta \leq \mu_b\} \\ (1-\pi_2)\pi_1\Big]\ g_{1,2}(\epsilon_1, \epsilon_2; \rho)\ f(\eta)\ d\eta\ d\varepsilon_2\ d\varepsilon_1 \tag{1.28}$$

Under assumptions (A1)-(A5), it holds that

$$\mathscr{P}_{(1,1)} + \mathscr{P}_{(0,0)} + \mathscr{P}_{(1,0)} + \mathscr{P}_{(0,1)} = 1$$

without invoke *coherency condition*.

1.4.2 Estimation of $\boldsymbol{\theta}$

All parameters are identify, then, using *maximum likelihood estimation*, we can construct the *likelihood function* as follows:

$$\mathscr{L}(Y,X,W,\theta) = \frac{1}{N}\sum_{i=1}^{N}\big[y_{i1}y_{i2}\ log(\mathscr{P}_{i\{1,1\}}) + (1-y_{i1})(1-y_{i2})\ log(\mathscr{P}_{i\{0,0\}}) \ + \\ + y_{i1}(1-y_{i2})\ log(\mathscr{P}_{i\{1,0\}}) + (1-y_{i1})y_{i2}\ log(\mathscr{P}_{i\{0,1\}})\big] \tag{1.29}$$

where $\mathscr{P}_{i\{j,k\}}$ is the probability to play $\{j,k\} \in (\{1,1\},\{0,0\},\{1,0\},\{0,1\})$ of the $i-th$ married couple.

Matrix of variances and covariances can be found using maximum likelihood estimation techniques (see Amemiya, 1985). Then

$$\Omega_{MLE} = -\left[E\frac{\partial^2\mathscr{L}(\mathbf{Y},\mathbf{X},\mathbf{W},\theta)}{\partial\theta\partial\theta'}\right]^{-1} \tag{1.30}$$

Which is known as Cramer-Rao lower bound.

1.5 Empirical application

1.5.1 Complete Information

Here, I assumed that husband and wife, they know each other complete when they interact strategically in order to decide if participate or not in the labor market.

1.5.2 Data base description

In order to characterize the strategic interaction and decisions of Mexican couples of spouses, particularly, those decisions which are related with the participation in the labor market, I used the "Encuesta Basal sobre el Ahorro, Crédito y Microfinanzas Rurales" (Basal Survey of Saves, Credit and Rural Micro-Finances), made by the "Banco del Ahorro Nacional y Servicios Financieros, Sociedad Nacional de Crédito, Institución de Banca de Desarrollo" (BANSEFI) and the "Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación" (SAGARPA), in 2004.

The entire survey was designed as a "natural experiment" in which it is attempt to capture changes and differences in social, economical and political terms, between those households in which unless one of its members belongs to one "Popular Credit and Save Society" (SACP, by its acronym in Spanish).

In this survey were interviewed 5767 households randomly selected: 2792 were "no treatment" and 2975 were "treatment"; distributed in 3 regions: North, Center and South[11]. Additionally, the households interviewed were divided in rural or urban communities. We discarded those couples in which one of its members is greater than 65 years old.

Table 1

Couples	N	North	Center	South
Mexico	3884	19.44%	38.59%	41.97%

$$Y_p = \begin{cases} 0 & \text{If } p \text{ don't participate in the labor market} \\ \\ 1 & \text{If } p \text{ participate in the labor market} \end{cases}$$

where $p \in \{1, 2\}$. Then, $(Y_1, Y_2) \in \{(1,1), (1,0), (0,1), (0,0)\}$. Outcomes of the games can be seen in Figure 7:

Figure 7

Outcomes of the Games

		PLAYER 2 (Wife) Y=1	Y=0
PLAYER 1	Y=1	907	2808
(Husband)	Y=0	42	127

[11]Region=1 (North): Aguascalientes, Baja California, Baja California Sur, Chihuahua, Coahuila, Durango, Nuevo León, San Luis Potos, Sinaloa, Sonora, Tamaulipas and Zacatecas; Region=2 (Center): Colima, Distrito Federal, State of Mexico, Guanajuato, Hidalgo, Jalisco, Michoacán, Morelos, Nayarit and Querétaro; Region=3 (South): Campeche, Chiapas, Guerrero, Oaxaca, Puebla, Quintana Roo, Tabasco, Tlaxcala, Veracruz and Yucatán.

Following Bjorn and Vuong (1985), I used social variables available in the survey which capture market participation decisions.

$$
\begin{aligned}
\mathbf{X}_1 &= \{AGEH, EDUCH, KIDS13\} \\
\mathbf{X}_2 &= \{AGEW, EDUCW, KIDS13\} \qquad (1.31) \\
\mathbf{W} &= \{AGEH^2, EDUCH\}
\end{aligned}
$$

$\mathbf{X}_p$ is known of the married couple, which reinforce the *complete informational* assumption. Covariates of all variables were summarized in the Table 2.

Table 2

Covariates of the Model

Variable	Description	Mean	Std. Dev.	Min	Max
AGEH	Age of the male patner	41.83	11.32	17	65
AGEW	Age of the female patner	38.62	10.89	14	65
EDUCH	Years of education (male patner)	7.16	4.52	0	17
EDUCW	Years of education (female patner)	6.81	4.28	0	17
KIDS13	Number of kids less than 14 years old	1.55	1.31	0	7
$AGEH^2$	Age Square of male patner (Experience proxy variable)	1878.49	976.075	289	4225

As can be seen, male partner average age is about 41 years old, female partner, is about 38. Additionally, years of education is about 7 years in both cases. In average, Mexican couples have 1.55 kids less than 14 years old. It was included the variable $AGEH^2$ as a proxy of the male experience. Without lose of generality, I assumed that couples "flip a coin" when they choose mixed strategies, so $\Pi_1 = 0.5$ and $\Pi_2 = 0.5$. Main results are presented in the next section.

1.5.3 Main Results

Table 3

(Standard Errors in parentheses)

Variable	**Player 1** (Husband)	**Player 2** (Wife)
$CONS_p$	4.6126*	-1.6807*
	(0.4985)	(0.2776)
AGE_p	-0.0385*	0.0125*
	(0.0083)	(0.0045)
$EDUC_p$	-0.0125	0.1894*
	(0.0172)	(0.0101)
$KIDS13_P$	0.0572	-0.1386*
	(0.0644)	(0.0378)
α_p	13.5730	-1.2276*
	(27.6374)	(0.1324)
ρ	0.5	

(*) Statistically significant at a 5% level.

As an additional assumption, ρ was picked up at 0.5, which means that there is a positive relationship between $(\varepsilon_1, \varepsilon_2)$. There were no significant changes with other values of ρ.

Table 4

Linear Index

(Standard Errors in parentheses)

Variable	$\boldsymbol{W}$
AGE_p^2	-0.1000
	(0.2169)
$EDUC_p$	-0.0100
	(0.0344)
μ_a	-0.0054
	(0.0102)
μ_b	-0.0048
	(0.0342)

1.5.4 Analysis of results

In both cases AGE_p's parameter were significant; nonetheless, husband's age was negative, which means that the more husband's age the less participation in the labor market, which is reasonable. On the other hand, female parameter is positive which means that wives have more incentives to participate in the labor market as they are older. Social and economic considerations can determine wives to participate in the labor market; for example, the more age that they have, children becomes "independent", and they have "free hands" to decide to work.

Education is not significative in the husband's case, but in the wife's case it is and positive. This means that the more years of education, the more incentives to participate in the labor market. In recent years, Mexican women have had an active roll in the labor market; specifically in those cases in which women are more prepared academically because they could get better jobs.

In the spouses decision to participate or not in the labor market, only wives care about the children. $KIDS13_p$ has negative coefficient, as predicted (Bjorn and Vuong (1984) obtain the same qualitative result in the female case). This means that children are an objective restriction when couple make their decisions to participate or not in the labor market. In Mexico, traditionally man is less worry about children in general which is directly related with gender culture in this country.

Interaction strategic parameter is very important in this analysis. As we can see in Table 2, this coefficient were significant only in the wives' case. This sign was negative, and means that they care about here husbands decisions. More over, she will not participate in the labor market if his husband decide to participate (in average). The fact that husband's parameter were not significant means that they don't care about their wives decisions (in average) to participate in the labor market.

All the parameters of the selection mechanism were not significant, which means that they don't follow this particular mechanism of equilibrium selection. Nonetheless, it is interesting that eventhough parameters of $EDUC_1$ were not significant, they both were negative (See Table 2 and 3).

1.6 Contrafactual Exercise

This section was included as a "contrafactual" ("what if") exercise, it is just hypothetical. Here is made the assumption that wife doesn't communicate his husband her decision to participate in the labor market, and viceversa. Under this assumptions, there is some interesting results.

1.6.1 Incomplete information

We will incorporate the assumption of *incomplete information*, seen above. This implies that players don't know payoff function of the other one, they should guess about it in order to take their decisions. In this context, we assume that players do not know Y_{-p}, instead of that, they incorporate beliefs in their payoff function as follows

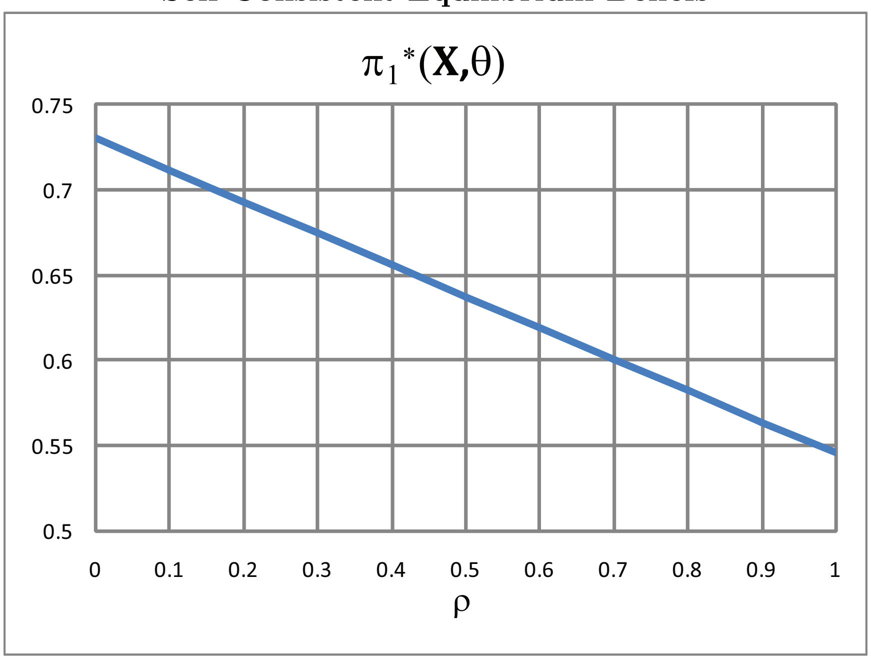
$$\pi_{-p} = Pr_p(Y_{-p} = 1 | Y_p = 1, X)$$

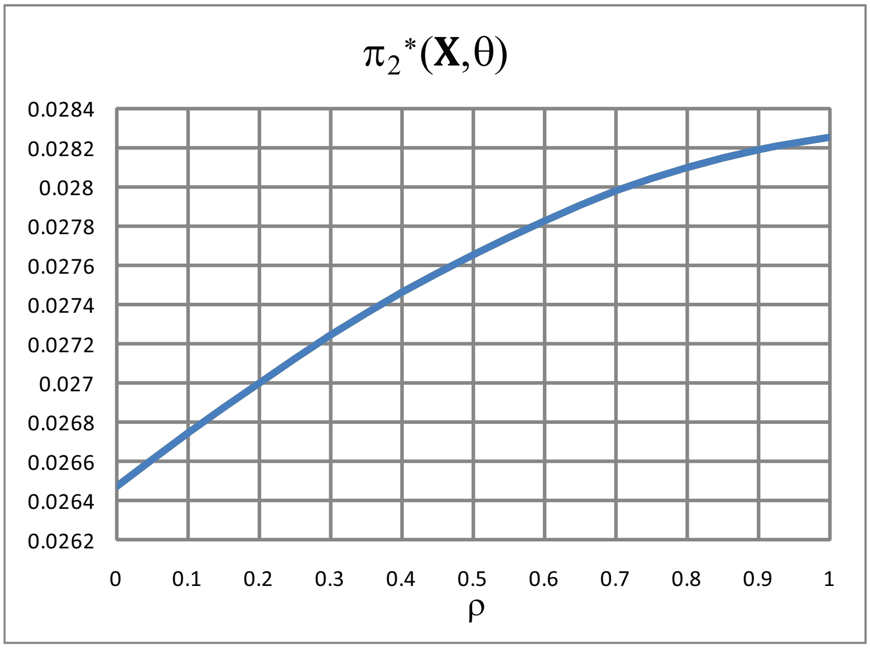
Which is the player p belief.

Using the same data based described above, $(Y_n, X_n)_{n=1}^{1962}$, and, for a given values of $\boldsymbol{\theta}$ and using a grid of ρ's values, we solve (1.11) and found the values of $\hat{\pi}_1(\boldsymbol{\theta}, \mathbf{x}), \hat{\pi}_2(\boldsymbol{\theta}, \mathbf{x})$, this values characterize behavior of a the set of *self-consistent equilibrium beliefs*. Graphically, can be seen in Figure 8:

Figure 8

Self-Consistent Equilibrium Beliefs





Given the fact that at different values of ρ, the value of $\hat{\boldsymbol{\theta}}$ didn't change qualitatively, then, without lost of generality, we picked-up $\rho = 0.5$. Values of $\hat{\pi}_1(\boldsymbol{\theta}, \mathbf{x}), \hat{\pi}_2(\boldsymbol{\theta}, \mathbf{x})$ are presented in Table 2.

Table 5

	$\hat{\pi}_1(\boldsymbol{\theta}, \mathbf{x})$	$\hat{\pi}_2(\boldsymbol{\theta}, \mathbf{x})$
$\rho = 0.5$	0.6375	0.0277
Biprobit	0.9438	0.2517

This results is interesting by itself, and help us to answer the question if there exists some sociocultural items that lead decisions of participation in the labor market of married people in Mexico. Husband's believes denote they guess that the probability of his wife participates in the market is near to zero when he decides to participate; on the contrary, wife's believes are greater than 0.5, which means that she believes that her husband will participate in the labor market when she decides to participate.

More over, self-consistent equilibrium beliefs, i.e., the probability of husband's participation, given that his wife participates in the labor market, is a decreasing function of ρ, the correlation between private information variables $(\varepsilon_1, \varepsilon_2)$; but is increasing in the wife's participation case given that husband participate.

Having in hand with this values, we solved (1.17). Results, are compared with *bivariate probit model* (approximation to the complete information model) in Table 3.

Table 6

Estimations Results

(Standard Errors in parentheses)

	Player 1 *Husband*	**Player 2** *Wife*	**Player 1** *Husband*	**Player 2** *Wife*
Variable	**Bivariate Probit**		**Incomplete**	
AGE_p	-0.0218*	0.0091*	-0.0488*	0.0145*
	(0.0047)	(0.0032)	(0.0113)	(0.0059)
$EDUC_p$	-0.0009	0.1091*	0.0033	0.1935*
	(0.0109)	(0.0087)	(0.0276)	(0.0144)
$REGI_p$	0.00808	0.1755*	0.0241	0.3167*
	(0.0691)	(0.0459)	(0.1669)	(0.0783)
$KIDS13_P$	N.A.	-0.0631*	N.A.	-0.1145*
		(0.0287)		(0.0503)
α_p	2.6033*	-2.1159*	187.8302*	-5.7305*
	(0.3039)	(0.2242)	(27.1241)	(0.6022)
ρ	-0.0538		0.5	
	(0.0695)		Picked up	

(*) Statistically significant at a 5% level.

1.6.2 Analysis of results

In general, both models behaves similarly. Only two variables observe one changing of sign and are significant in both models: AGEH and α_1. Then, this changes only are observed in the husband equation.

AGEH is negative and significant in the incomplete information model. This means that the more age of the husband, he will not participate in the labor market, which sounds logic. On the contrary, AGEW is positive and significant, which means that wives could participate in the labor market when they reach a mature age. In this case, they have less restrictions, for example, the attendance of their children.

In the incomplete information case, husband's equation shows that education, EDUCH, is not significant. Man has a high propensity to participate in the labor market despite his education; this can be explained as follows: in Mexico, gender discrimination or cultural conceptions could be driven decisions of husbands: "he should work, he is the man of the house, and education doesn't matter". Interestingly, EDUCW is positive and significant, which means that the more wives' education will participate in the market labor.

REGI is not significant in husband's equation and a similar analysis can be made as in the EDUCH's case, because cultural behavior. Husband will participate despite of the region. In the wives' case, things are different. In this case, the value of the parameter is positive and significant, this means that regional situations could have influences in the labor market participation. The positive sign, means that in regions 2 or 3 (center and south), there is more wives' participation in the market labor because of economic and sociocultural problems. North is richer and conservative (culturally speaking), than in the Center region.

As it was mention, KIDS13 only was included in wives' equation and was negative and significant. This result is rational: we hope that, the more children less than 14 years old, less wives' participation in the labor market.

As can be seen, in the *bivariate probit model* both parameters are negative and significant, enter-enter, $(Y_1 = 1, Y_2 = 1)$ hurts both players; nonetheless, in the *incomplete information model*, α_1 is positive and significant, and α_2 is negative and significant. This means that husband will participate if he guess that his wife will do it too; on the contrary, if wife guess that the husband will participate in the labor market, she will not do it, then, she is affected by her beliefs about the husband's actions. Then, we can conclude that husband and wife interact strategically and, at the same time, that there exists some sociocultural factors that lead behavior of married couples in Mexico.

1.7 Concluding remarks

Decisions to participate in the labor market can be modeled in the context of the game theory, and here is presented in the Mexican labor market. In the context of *complete information* and multiple equilibrium game, it was modeling specific mechanism of equilibrium selection. Results reveal that participation decisions in the labor market that come from this game structure have more influence over the wife than the husband. Husband's decisions are not essentially affected by wife's decisions; but, wife's decisions are directly affected by his husband's decisions ($\alpha_2 < 0$ and significant). Husband participation decisions are only affected by the age; wife is affected by age, education and kids less than 14 years old. Finally, all variables from selection mechanism were not significant.

This results are very important but there is more to do. For example, mechanism of selection could be analyzed in the context of other games with described characteristics here. Main problem is to get information in order to use this model. Additionally, simulations can be done in order to better understand under which conditions works correctly the designed mechanism of equilibrium selection.

Chapter 2

Semi-Empirical Likelihood Estimation of Manufacturing Interaction-Based Model with Asymmetric Information

2.1 Interaction-Based Models

Let us illustrate ideas by considering the following example (similar to the last chapter): suppose we have a 2 x 2 game in which players must simultaneously (i.e., before observing their opponent's choice) choose between two actions: "Enter" or "Don't Enter". We can assume the following payoff matrix, without loss of generality:

Figure 1

A simple 2 x 2 game

		PLAYER 2	
		Enter	Don't
PLAYER 1	Enter	$t_1 - \pi_2\alpha_1, t_2 - \pi_1\alpha_2$	$t_1, 0$
	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 $\mathcal{P}(t)$.

π_1=Probability that player 1 chooses Enter.

π_2=Probability that player 2 chooses Enter.

Let

$E_1[\pi_2]$≡Player 1's belief that player 2 will choose Enter.

$E_2[\pi_1]$≡Player 2's belief that player 1 will choose Enter.

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

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

Given the fact that the payoff of "don't enter" has been normalized to zero in this case, then players 1 and 2 will choose "enter" if and only if $E_1[u_1^{Enter}] > 0$ and $E_2[u_2^{Enter}] > 0$. This implies that Bayesian-Nash equilibrium beliefs must satisfy

$$E_1[\pi_2] = 1 - \mathcal{P}(E_2[\pi_1]\alpha_1) \textit{ and } E_2[\pi_1] = 1 - \mathcal{P}(E_1[\pi_2]\alpha_2) \tag{2.1}$$

Now, in order to include econometric considerations for estimating beliefs, we need to consider the stochastic characteristics of our game. Payoff functions are unobservable. Suppose t_1 and t_2 can be expressed as functions of $(\mathbf{X}_1, \varepsilon_1)$ and $(\mathbf{X}_2, \varepsilon_2)$ respectively. The following assumptions preserve the stochastic and informational assumptions of this game.

A1.-$\mathbf{X}_1 \in \mathbb{R}^k$ and $\mathbf{X}_2 \in \mathbb{R}^k$ 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 $\mathbf{X}_p$ for $p \in \{1, 2\}$

A4.- *At the time the game is played*, the realizations of $(\mathbf{X}_1, \varepsilon_1)$ and $(\mathbf{X}_2, \varepsilon_2)$ 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 $(\mathbf{X}_p, \varepsilon_p)$, $p \in \{1, 2\}$.

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

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

Suppose that, without loss of generality, we can parameterize private information, t_1 and t_2, in the next way:

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

where the parameter vector $\boldsymbol{\beta}$ is known by both players, and is assumed to be the same. Then, Bayesian-Nash equilibrium conditions become:

$$\begin{aligned} E_2[\pi_1] &= \int_{\boldsymbol{x}} G(\boldsymbol{\beta'}\mathbf{X}_1 - E_1[\pi_2]\alpha_1) dF(\boldsymbol{x}) \\ E_1[\pi_2] &= \int_{\boldsymbol{x}} G(\boldsymbol{\beta'}\mathbf{X}_2 - E_2[\pi_1]\alpha_2) dF(\boldsymbol{x}) \end{aligned} \tag{2.2}$$

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

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

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

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

B4.-The distribution $G(\epsilon)$ 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 of $F(\boldsymbol{x})$. We 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 in particular it can be applied to models in which all agents can belong to one of a *finite* number of "types", and each type is public information, which is our case. Player's types contain some information about their private payoffs. This would be the case for example if in the model presented above 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 $\mathcal{X}_1 \cup ... \cup \mathcal{X}_T = \mathbb{R}^k$, or, which is the same, $\mathcal{X}_1 \sqcup ... \sqcup \mathcal{X}_T = \mathbb{R}^k$ ($\sqcup$, means *disjoint* union). We say that player p belongs to *type* τ_t if and only if $\boldsymbol{X}_p \in \mathcal{X}_t$.

Then, for all possible applications, the purpose is to estimate *simultaneously* the following elements of the model:

1.- The structural 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 above description)

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

4.- 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 under the assumption that observed outcomes are the result of Bayesian-Nash equilibria. The link between all these beliefs is given by the corresponding equilibrium restrictions that they must satisfy (equation (2.2) in the example presented above). The issues of existence and uniqueness of an equilibrium are crucial and they will be addressed later, along with the asymptotic properties of the proposed model in this paper.

2.2 Brief overview of empirical likelihood (EL)

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 $E[Z] = \mu$. The optimal weights would be the solution of the problem

$$\max_{\{p_i\}_{i=1}^N} \sum_{i=1}^N \log p_i \quad \text{subject to: } p_i > 0, \ \sum_{i=1}^N p_i = 1 \text{ and } \sum_{i=1}^N p_i Z_i = \mu$$

That is, to maximize the empirical log-likelihood $\sum_{i=1}^N p_i$ subject to the weights being a well-behaved pdf, and the data obeying E[Z]=μ 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. This 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 $\{p_i\}_{i=1}^N$ 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 interval could be based on the statistic $\mathcal{R}(\mu_0)$. The α-level confidence interval, for example, would be constructed as the set of $\mu \in \mathbb{R}$ such that $\mathcal{R}(\mu) \leq c_\alpha = 1 - \alpha$. 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 $\hat{p}_i = 1/N$.

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 $\mathbf{Z} \in \mathbb{R}^d$ there is a parameter $\boldsymbol{\theta} \in \mathbb{R}^p$ and a vector valued function $m(\mathbf{Z}, \boldsymbol{\theta}) \in \mathbb{R}^s$ such that $E[m(\mathbf{Z}, \boldsymbol{\theta})] = 0$. For a fixed $\boldsymbol{\theta}$, the corresponding EL problem is to solve

$$\max_{\{p_i\}_{i=1}^N} \sum_{i=1}^N \log p_i \quad \text{subject to: } p_i > 0,\ \sum_{i=1}^N p_i = 1 \text{ and } \sum_{i=1}^N p_i m(\mathbf{Z}, \boldsymbol{\theta}) = 0.$$

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

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

where q is the rank of $Var[m(\mathbf{Z}, \boldsymbol{\theta}_0)]$. Confidence regions can be built and hypothesis 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 $\hat{\boldsymbol{\theta}}$ is simply given by the solution of $\sum_{i=1}^N m(\mathbf{z}_i, \hat{\boldsymbol{\theta}}) = 0$ and the resulting optimal weights are the uniform ones, $\hat{p}_i = 1/N$. 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 $\mathbf{Z} \in \mathbb{R}^k$ is assumed to have a known parametric functional form given by $f(y|\mathbf{Z}, \boldsymbol{\theta})$, but that the marginal pdf of $\mathbf{Z}$ is unknown and denoted by $dF(\mathbf{z})$. The joint pdf of $(Y, \mathbf{Z})$ would then be given by $f(y|\mathbf{z}, \boldsymbol{\theta})dF(\mathbf{z})$. Suppose now that we know that $E[\psi(\mathbf{Z}, \boldsymbol{\theta})] = 0$ for some function $\psi \in \mathbb{R}^p$. EL would estimate $\boldsymbol{\theta}$ and $\{p_i\}_{i=1}^N$ by solving

$$\max_{\boldsymbol{\theta},\{p_i\}_{i=1}^N} \sum_{i=1}^{N} log\ f(y_i|\boldsymbol{z}_i,\boldsymbol{\theta}) + \sum_{i=1}^{N}\ log\ p_i$$

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

Qin (1994, 2000) called the combination of 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 $\mathbf{z}_1, ..., \mathbf{z}_n$ is directly observed form a distribution $F(\mathbf{z})$ while another sample $\mathbf{x}_1, ..., \mathbf{x}_n$ has density $\int p(\mathbf{x}|\mathbf{z})dF(\mathbf{z})$ where $p(\mathbf{x}|\mathbf{z})$ is parameterized as $p(\mathbf{x}|\mathbf{z},\boldsymbol{\theta})$. Parametric and empirical likelihoods have also been combined in Bayesian models. Lazar (2000) analyzed the product of prior density on the univariate mean and an empirical likelihood for that mean.

Kitamura (2006) has a comprehensive summary of Empirical Likelihood techniques in which he studies some computational strategies in order to solve the problems studied above. The methodology proposed here is a particular case of semi-empirical likelihood estimation.

2.2.1 Empirical Likelihood and GMM

Every GMM problem can also be estimated using EL. Asymptotic equivalence to first order approximation between GMM and EL has been well documented in a variety of settings (Owen (2001) and Kitamura (2006) are the best comprehensive references). It has also been established that EL improves on the small sample properties of GMM. However, other closely estimators that improve on the small sample properties of GMM have also been developed: continuous updating (CUE) -also called "Euclidian Likelihood" by Owen (2001)- and exponential titling estimators (ET). All these belong to a class of Generalized Empirical Likelihood (GEL) estimators. To first order of approximation, 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 10 times respectively. He also found that EL's power ranking was best for hypothesis further from the null.
The most relevant results to the problem we address here are 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.

2.3 Investment Strategy 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 simultaneously by n different players . The application presented here corresponds to the latter case.

The 2×2 game describe above was used to illustrate the properties of the proposed empirical likelihood estimator. A brief description of an investment strategy 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 investment decision in an environment of asymmetric information. We will now define the meaning of "investment decision" by specifying the particular space of actions for this model.

2.3.1 The model

For firm i let's denote: $d_i \equiv$ Firm's industry, and $k_i \equiv$ Firm's technological category. Note that $d_i \in \{1, 2, \ldots, D\}$, and $k_i \in \{LT, SS, SL, HT\}$[1].

Timing of firm's decisions

At time t the firm must choose to increase or not investment in period *t+1*. All the firms make this decision simultaneously (i.e., before observing what other firms have optimally chosen to do) and in the context of *asymmetric information* which will be described bellow. Let's denote *ith* firm's decisions as follows:

$$Y(i) = \begin{cases} 1 & \text{If firm is } \textit{passive.} \\ \\ 2 & \text{If firm is } \textit{neutral.} \\ \\ 3 & \text{If firm is } \textit{aggressive.} \end{cases}$$

How firms can be affected by others' decisions is explained next.

2.3.2 Strategic interaction among firms

In economics investment is defined as the act of incurring an immediate cost in the expectation of future rewards, Dixit and Pindyck (1994). Given the fact that the investment is relatively irreversible, and there is uncertainty to obtain the future expected reward, we expect that firms care about the others' actions in their own

[1]LT≡low-tech segment, SS≡stable tech-short horizon segment, SL≡stable tech-long horizon segment, and HT≡hi-tech segment. Hall and Vopel (1997)

decisions because they could increase (or reduce) the probability of failure in the expected reward of the investment, seen as a sunk cost. Firms interact in many dimensions, but because a firm's relative size in its industry has been consistently cited as a determinant of investment, as well as market structure, the present model will attempt to *analyze how small, medium and large firms interact.* The goal is to answer the following questions:

1.- Do small firms care about the investment decision made by other small firms? Do they care about the decisions made by medium and large firms?

2.-Do medium firms care about the investment decision made by other medium firms? Do they care about the decisions made by small and large firms?

3.-Do large firms care about the investment decision made by other large firms? Do they care about the decisions made by small and medium firms?

Choice rules will be modeled in such a way that allows us to test separately the influence of other firms' investment on a specific firm's investment decision. In particular, without loss of generality, we will model how the representative firm (small, medium or large) could be affected if the others (small, medium or large) decide to be "aggressive" in the investment sense.

2.3.3 Decision rules

We assume that the decision made by the firms are ordered according to some criteria. Let be the "types", $k = \{S, M, L\}$, if firm is "small", "medium", and "large" respectively:

$$u_i^k = \alpha^S \pi_A^S + \alpha^M \pi_A^M + \alpha^L \pi_A^L + \boldsymbol{\beta}' X_i + \varepsilon_i \tag{2.3}$$

Were π_A^k is the proportion of the population of size "k" that will choose to be aggressive. The remaining variables, $\boldsymbol{X}$ and ε, will be described below.

Optimal decision rules

Given the ordered nature of our model we can define: let be $\zeta_1 < \zeta_2$ "threshold" parameters (not observed by the econometrician), such that:

$$Y(i) = \begin{cases} 1\ (passive) & if\ u_i^k \leq \zeta_1 \\ \\ 2\ (neutral) & if\ \zeta_1 < u_i^k \leq \zeta_2 \\ \\ 3\ (aggressive) & if\ u_i^k > \zeta_2 \end{cases}$$

Due to the asymmetric information nature of the model, the proportions π_A^S, π_A^M, and π_A^L are not public information. Firms will then maximize the expected version of the payoff function (2.3). This shall be carefully detailed below.

2.3.4 Strategic interaction

Interaction coefficients

As it was mentioned above, the goal of the model is to estimate the influence of the other firms' choices on an individual (representative) firm investment decision. For a firm of size $k = \{S, M, L\}$, α_A^S, α_A^M, and α_A^L indicates the influence of population of small, medium, and large firms investment decisions respectively, on the firms own investment choice.

Why would firms interact?

Modern models of firm survival argue that a firm's innovation capabilities, which are influenced, for example, by the investment in R&D, determine its chances of surviving in the long run. It is reasonable, then, to think that firms would interact based in long-run consequences of investment.

2.3.5 Determinants of investment

Tobin's Q compares the capitalized value of a marginal investment in real capital to its replacement cost. According to the net present value (NPV) theory of investment, Aradillas-Lopez (2007), the firms should adjust its investments decisions according to changes in Q_i. Then, we used $\Delta Q_i = Q_{i,t} - Q_{i,t-1}$: level change in Tobin's Q from *t-1* to t, as explanatory variable. In order to capture the short-term firm's performance we used the percentage change of its sales $\Delta\%S_i = \frac{(S_{i,t}-S_{i,t-1})}{S_{i,t-1}}$. How firms acted in the past, could influence the future decision, that is why the lag of the decision variable, $y_{i,t-1}$, was included: to be passive, neutral or aggressive, in the past period.

2.3.6 Distributional assumptions

Let $\boldsymbol{X} \equiv \{y_{i,t-1}, \Delta\%S_i, \Delta Q_i\}$. Then, we will assume the following:

i.-$\boldsymbol{X}$ has a unknown joint cdf given by $G_{\boldsymbol{X}}(\boldsymbol{x})$. Whose pdf is denoted as $dG_{\boldsymbol{X}}(\boldsymbol{x})$.

ii.- Conditional of $\boldsymbol{X}$, ε, has a marginal cdf given by $F(\epsilon)$. We will assume a particular functional form for this distribution with parameters independent of $\boldsymbol{X}$.

2.3.7 Informational assumptions

We will make the following assumptions regarding the information structure of the model:

i.- When firms make their optimal choices, variables $\boldsymbol{X}$ and ε, are privately known.

ii.- Variables $\boldsymbol{X}$ and ε become available (for the econometrician) some time after the optimal choices have been made. The variable ε, remains unknown to the econometrician.

2.3.8 Beliefs and equilibrium conditions

As we mentioned above, when making their optimal choices, firms can't observe the population proportions of π_A^S, π_A^M, and π_A^L. Firms will then maximize the expectation in their payoff function (2.3). Let

$$\begin{aligned} E_i[\pi_A^S] &= \text{Firm i's expectation of } \pi_A^S \\ E_i[\pi_A^M] &= \text{Firm i's expectation of } \pi_A^M \\ E_i[\pi_A^L] &= \text{Firm i's expectation of } \pi_A^L \end{aligned} \tag{2.4}$$

In equilibrium, due to the informational assumptions of the model, all firms must have the same beliefs. Let's denote these beliefs as $\bar{\pi}_A^S$, $\bar{\pi}_A^M$, and $\bar{\pi}_A^L$. Linearity of the payoff function (2.3) allows to simply plug in these beliefs instead of the true probabilities in order to compute expected payoffs, which are described as follows.

$$\bar{u}_i^k = \alpha^S \bar{\pi}_A^S + \alpha^M \bar{\pi}_A^M + \alpha^L \bar{\pi}_A^L + \boldsymbol{\beta}' \boldsymbol{X}_i + \varepsilon_i \tag{2.5}$$

Then, decisions of the firms will be driven by:

$$Y(i) = \begin{cases} 1\ (passive) & if\ \bar{u}_i^k \leq \zeta_1 \\ \\ 2\ (neutral) & if\ \zeta_1 < \bar{u}_i^k \leq \zeta_2 \\ \\ 3\ (aggressive) & if\ \bar{u}_i^k > \zeta_2 \end{cases} \tag{2.6}$$

2.3.9 Estimation and results

Identification

Identification concerns are very important in interaction-based models. This section examines issues related to the proposed model. Let's denote:

$$\begin{aligned}
\boldsymbol{\theta}_1 &= (\bar{\pi}_A^S, \bar{\pi}_A^M, \bar{\pi}_A^L) \\
\boldsymbol{\theta}_2 &= (\alpha^S, \alpha^M, \alpha^L, \zeta_1, \zeta_2, \boldsymbol{\beta})' \qquad (2.7) \\
\boldsymbol{\theta} &= (\boldsymbol{\theta}_1', \boldsymbol{\theta}_2')'
\end{aligned}$$

Now let

$$\delta(\boldsymbol{\theta}, \boldsymbol{X}) = \alpha^S \bar{\pi}_A^S + \alpha^M \bar{\pi}_A^M + \alpha^L \bar{\pi}_A^L + \boldsymbol{\beta}' \boldsymbol{X} \qquad (2.8)$$

Conditional of $\boldsymbol{X}$ and $\overline{\boldsymbol{\pi}}_A = \{\bar{\pi}_A^S, \bar{\pi}_A^M, \bar{\pi}_A^L\}$, we have the following results:

$$\begin{aligned}
Pr(passive|\boldsymbol{X}, \overline{\boldsymbol{\pi}}_A) &= Pr(\boldsymbol{Y} = 1|\boldsymbol{X}, \overline{\boldsymbol{\pi}}_A) = \\
&= Pr(\bar{u}^k \leq \zeta_1) = \\
&= F(\zeta_1 - \delta(\boldsymbol{\theta}, \boldsymbol{X})) \\
Pr(neutral|\boldsymbol{X}, \overline{\boldsymbol{\pi}}_A) &= Pr(\boldsymbol{Y} = 2|\boldsymbol{X}, \overline{\boldsymbol{\pi}}_A) = \\
&= Pr(\zeta_1 < \bar{u}^k \leq \zeta_2) = \\
&= F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X})) - F(\zeta_1 - \delta(\boldsymbol{\theta}, \boldsymbol{X})) \\
Pr(aggressive|\boldsymbol{X}, \overline{\boldsymbol{\pi}}_A) &= Pr(\boldsymbol{Y} = 3|\boldsymbol{X}, \overline{\boldsymbol{\pi}}_A) = \\
&= Pr(\bar{u}^k > \zeta_2) = \\
&= 1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))
\end{aligned}$$

Where $F(\bullet)$ is the cdf of ε. Using last definition, beliefs can be modeled as follows: let be $K = \{S, M, L\}$

$$\begin{aligned}
\bar{\pi}_A^k &= Pr(aggressive|\boldsymbol{X}, k) = \\
&= Pr(Y = 3|\boldsymbol{X}, k) = \\
&= E_{k=K}[Pr(Y = 3|\boldsymbol{X}, k)|k = K] = \\
&= \frac{\sum_{i=1}^{N} [1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))] \mathbf{1}\{k = K\}}{\sum_{i=1}^{N} \mathbf{1}\{k = K\}} \qquad (2.9)
\end{aligned}$$

Define:

$$
\begin{aligned}
\psi_1(\boldsymbol{\theta}, \boldsymbol{X}) &\equiv \bar{\pi}_A^S - \frac{\sum_{i=1}^N [1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))]\mathbf{1}\{k = S\}}{\sum_{i=1}^N \mathbf{1}\{k = S\}} \\
\psi_2(\boldsymbol{\theta}, \boldsymbol{X}) &\equiv \bar{\pi}_A^M - \frac{\sum_{i=1}^N [1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))]\mathbf{1}\{k = M\}}{\sum_{i=1}^N \mathbf{1}\{k = M\}} \\
\psi_3(\boldsymbol{\theta}, \boldsymbol{X}) &\equiv \bar{\pi}_A^L - \frac{\sum_{i=1}^N [1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))]\mathbf{1}\{k = L\}}{\sum_{i=1}^N \mathbf{1}\{k = L\}} \\
\Psi(\boldsymbol{\theta}, \boldsymbol{X}) &\equiv \Big(\psi_1(\boldsymbol{\theta}, \boldsymbol{X}), \psi_2(\boldsymbol{\theta}, \boldsymbol{X}), \psi_3(\boldsymbol{\theta}, \boldsymbol{X})\Big)' \qquad (2.10)
\end{aligned}
$$

Then, Bayesian-Nash Equilibrium beliefs must satisfy

$$
\int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}, X) dG_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbf{0} \qquad (2.11)
$$

Existence of equilibria

For a given value of $\boldsymbol{\theta}_2$, we're interested in knowing if exists a set of beliefs $\boldsymbol{\theta}_1$ such that the equilibrium condition (2.11) is satisfied. A sufficient condition for the existence of equilibria is that the marginal distribution of ε be continuous. Existence of equilibria for an arbitrary value of $\boldsymbol{\theta}_2$ follows from Brouwer'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 equilibria

Uniqueness is a very important question. If, for the true values of $\boldsymbol{\theta}_2$ exist more than one set of beliefs $\boldsymbol{\theta}_1$ that satisfy equilibrium condition (2.11), 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. The question of uniqueness can be analyzed by looking at the Jacobian

$$\nabla_{\boldsymbol{\theta}_1} \int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0, \boldsymbol{x}) dG_{\boldsymbol{X}}(\boldsymbol{x}) \tag{2.12}$$

Where as before $\boldsymbol{\theta}_2^0$ represents the true population values of $\boldsymbol{\theta}_2$. Local unique equilibrium will be guaranteed if the Jacobian

$$\nabla_{\boldsymbol{\theta}_1} \int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}_1^*, \boldsymbol{\theta}_2^0, \boldsymbol{x}) dG_{\boldsymbol{X}}(\boldsymbol{x})$$

has a rank equal to three -full rank condition-, where $\boldsymbol{\theta}_1^*$ is a solution of

$$\int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0, \boldsymbol{x}) dG_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbf{0}$$

A sufficient condition for global uniqueness would be to assume that the Jacobian $\nabla_{\boldsymbol{\theta}_1} \int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0, \boldsymbol{x}) dG_{\boldsymbol{X}}(\boldsymbol{x})$ has either: (i) only strictly positive principal minors or (ii) only strictly negative principal minors for all $\boldsymbol{\theta}_1 \in [0,1]^3$. This is a version of Gale-Nikaido theorem that guarantees that $\int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2^0, \boldsymbol{x}) dG_{\boldsymbol{X}}(\boldsymbol{x})$ is a one-to-one function of $\boldsymbol{\theta}_1$ and therefore, that the equilibrium is unique. Summarizing, it says that the Jacobian not only has to be non-singular, but is also has to remain either positive quasi definite or negative quasi definite for all values of $\boldsymbol{\theta}_1$.

Existence and uniqueness have to do with identification of $\boldsymbol{\theta}_1$, the vector of beliefs. The functional form assumed for the expected payoff function requires two additional conditions for the identification of $\boldsymbol{\theta}_2$. These conditions are necessary for the asymptotic invertibility of the Hessian for the first order conditions satisfied by the EL estimator.[2]

i.- All equilibrium beliefs $\boldsymbol{\theta}_1^0$ must be strictly between zero and one. That is, in equilibrium the population probability of choosing the action to be aggressive must be strictly positive and this must hold for all type of firms (S,M,L). This is a necessary condition for identification of $\boldsymbol{\theta}_2$.

[2] The role played by these identification condition is parallel to the one played by the conditions necessary to assume invertibility of the information matrix in the usual MLE problems.

ii.- The conditional distributions $G(\boldsymbol{X}|k=S)$, $G(\boldsymbol{X}|k=M)$ and $G(\boldsymbol{X}|k=L)$ are not identical. 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 be neutral, for example. If (ii) is violated, it would imply that there is no strategic interaction that takes place in the type dimension (size): there is nothing essentially different between small and medium firms, etc. Violations to (i) or (ii) seem implausible to reality.

Estimation

Conditional likelihood

Having dealt with identification, we now present the estimator. The log-likelihood function of Y given $\boldsymbol{X}$ is given by:

$$\begin{aligned} log\ f(Y|\boldsymbol{X},\boldsymbol{\theta}) &= \mathbf{1}\{Y=1\}log[F(\zeta_1-\delta(\boldsymbol{X},\boldsymbol{\theta}))]+ \\ &+ \mathbf{1}\{Y=2\}log[F(\zeta_2-\delta(\boldsymbol{X},\boldsymbol{\theta}))-F(\zeta_1-\delta(\boldsymbol{X},\boldsymbol{\theta}))]+ \\ &+ \mathbf{1}\{Y=3\}log[1-F(\zeta_2-\delta(\boldsymbol{X},\boldsymbol{\theta}))] \end{aligned} \tag{2.13}$$

Where $F(\bullet)$ is the ε's cdf.

Empirical Likelihood estimator

The proposed Empirical Likelihood estimator $\hat{\boldsymbol{\theta}}^{EL}$ is the solution to:

$$\max_{\boldsymbol{\theta},\{p_i\}_{i=1}^N} \sum_{i=1}^N log\ f(y_i|\boldsymbol{x}_i,\boldsymbol{\theta})+\sum_{i=1}^N log\ p_i \tag{2.14}$$

subject to

$$p_i \geq 0, \ \sum_{i=1}^{N} p_i = 1, \ \sum_{i=1}^{N} p_i \Psi(\boldsymbol{x}_i, \boldsymbol{\theta}) = 0 \tag{2.15}$$

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

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

$$E[\nabla_{\boldsymbol{\theta}} log \ f(\boldsymbol{Y}|\boldsymbol{X}, \boldsymbol{\theta}_0)] \ and \ E[\Psi(\boldsymbol{X}, \boldsymbol{\theta}_0)]$$

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

$$\max_{\boldsymbol{\theta}} \ \sum_{i=1}^{N} log \ f(\boldsymbol{y}_i|\boldsymbol{x}_i, \boldsymbol{\theta}) \text{ subject to } \ \frac{1}{N}\sum_{i=1}^{N} \Psi(\boldsymbol{x}_i, \boldsymbol{\theta}) = 0$$

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 $dG(\boldsymbol{X})$ and the parameters of interest.

3.- Using available additional information about the population distribution of $\boldsymbol{X}$ increases the efficiency of $\hat{\boldsymbol{\theta}}^{EL}$.

2.4 Empirical Application

We are trying to know how firms interact with each other given their "type" (size) and their "actions" (be aggressive, neutral or passive), and considering that they belong to a particular "industry". The United States manufacturing sector in which an "industry" is defined by the SIC classification code (Standard Industrial Classification), was studied.[3]. All information was collected from Standard and Poor's Industrial

[3]SIC has been replaced by the North American Industrial Classification System (NAICS) codes, which identify companies according to economic, subsector and industry groups. There is a close link between them.

Compustat-North Amercia data set.

Using Compustat we identified 9 of the most numerous industries whose SIC number were {2834, 2836, 3674,3845, 2911, 3089, 3312, 3559, 3714}. They belong to the tech segments 1, 2 and 3, accordingly with Hall and Vopel (1997), who proposed a classification table for 4-digit SIC industries based on Chandler's technological segments (see appendix for details).

Time period considered here was t={1991, 1993, 1995}. The difference between years is attained in attempt to mitigate the effect of time-dependance. Each industry and every year were treated as a cross-section, then, all observations were pooled together, resulting a sample size of 986[4]. Let PISHIP and PIINV denote the industry-annual price deflators for the value of shipments and total capital expenditures respectively taken from the NBER-CES Manufacturing Database.

Decision variable in our model, $y_{i_{st}}$, was constructed as follows[5]: we used the rate capital investment, $R_{i_{st}}$, which is equal to $I_{i_{st}}/K_{i_{st-1}}$ where:

$I_{i_{st}}$: net capital investment by firm i. Capital expenditures in property plant and equipment (Compustat item: data30) deflated by PIINV.

$K_{i_{st-1}}$: net capital stock made by firm i at the end of the period t-1. It was measured as the net value of property, plant and equipment (Compustat item: data8)[6], deflated by an annual capital stock deflator which was constructed for each industry using PIINV starting in 1958 and ending in 1997[7].

[4]Similar exercises were made for each tech-segment, but results essentially didn't change compared with the data which were pooled.

[5]the subscript i_{st} means: the firm i, which belongs to the SIC "s" at year "t"

[6]This data item is defined in Compustat as "the cost of tangible fixed property used in the production of revenue, less accumulated depreciation".

[7]It was necessary to use a linear projection with the purpose of obtain index's value for the year 1997, assuming constant depreciation rate across all industries.

Criteria used for the main variable, $y_{i_{st}}$, was:

$$y_{i_{st}} = \begin{cases} 1 & \text{If } R_{i_{st+1}} \leq 0.75 * R_{i_{st}} \text{ (passive).} \\ 2 & \text{If } 0.75 * R_{i_{st}} < R_{i_{st+1}} \leq 1.25 * R_{i_{st}} \text{ (neutral).} \\ 3 & \text{If } R_{i_{st+1}} > 1.25 * R_{i_{st}} \text{ (aggressive).} \end{cases}$$

It means that the firm is considered *passive* if its rate capital investment in the next period $(t+1)$ is 25% less or equal than the rate at current period (t). The firm is considered *aggressive* if the rate capital investment is 25% greater than the rate of the current period. A firm will be *neutral* if its rate capital investment in t+1 is something in between.

Then we have:

$$y_{i_{st}} = \begin{cases} 1 & \text{365 firms (37.71\%)} \ (passive) \\ 2 & \text{193 firms (19.94\%)} \ (neutral) \\ 3 & \text{410 firms (42.36\%)} \ (aggressive) \end{cases}$$

This tries to model the decisions taken by the firms which supposedly used (2.5) and (2.6) as action choice criteria. On the other hand, let be

$S_{i_{st}}$: firms' net sales (Compustat item: data12)[8] deflated by PISHIV. This variable will be used for constructing both: percent change in sales (explanatory variable) and "size" (type of the firms).

[8]It was used Employees (Compustat item: data29) as a criteria for determining $Size_{i_{st}}$, but results did not change essentially.

2.4.1 Size (Type of the firm)

Let's define $Size_{i_{st}} \equiv \frac{S_{i_{st}}}{median(S_{i_{st}})}$, as the size (type) of the *i's* firm. We have three "types": *small*, *medium* and *large*. Criteria which define types used in this paper were:

$$Size_{i_{st}} = \begin{cases} S & \text{if i's firm } S_{i_{st}} \leq 1/3 \text{ (Small).} \\ \\ M & \text{if i's firm } S_{i_{st}} \in (1/3, 2/3] \text{ (Medium).} \\ \\ L & \text{if i's firm } S_{i_{st}} > 2/3 \text{ (Large).} \end{cases}$$

The median, instead of the mean, was used with the purpose of getting away extremum values in the construction of the size index.

2.4.2 Explanatory Variables

Tobin's Q was calculated as in Jovanovic and Rousseau (2003).

$$Q_{i_{st}} = FMV_{i_{st}} / FBV_{i_{st}}$$

Where $FMV_{i_{st}}$ is the *Firm Market Value* which is the addition of current value of common equity (Compustat items: data24×data25), book value of preferred stock (Compustat item: data130) and short and long-term debt (Compustat items: data34×data9). And $FBV_{i_{st}}$ is the *Firm Book Value*, which is the sum of book value of common equity (Compustat item: data60), book value of preferred stock (Compustat item: data130) and short and long-term debt (Compustat items: data34×data9). $\Delta\%S_{i_{st}} = \frac{(S_{i_{st}} - S_{i_{st-1}})}{S_{i_{st-1}}}$, is the percentage change of firm's net sales, were $S_{i_{st}}$ was computed as above. $y_{i_{st-1}}$ was derived using the same criteria that defined $y_{i_{st}}$, but for a period before. Then, we have:

$$\boldsymbol{X}_{i_{st}} = (y_{i_{st-1}}, \Delta\%S_{i_{st}}, \Delta Q_{i_{st}})$$

2.4.3 Semi-Empirical Likelihood Estimation

Given the fact that we have 3 periods, 9 SIC industry and 3 types (S,M,L), there is 81 "moment condition", i.e., the vector defined in (2.8) here is 1 $\times$ 81.

$$\Psi(\boldsymbol{\theta}, \boldsymbol{X}) \equiv \Big(\psi_1(\boldsymbol{\theta}, \boldsymbol{X}), \ldots, \psi_{81}(\boldsymbol{\theta}, \boldsymbol{X})\Big)' \tag{2.16}$$

and should satisfy

$$\int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}, \boldsymbol{X}) dG_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbf{0} \tag{2.17}$$

ε-distribution

We assume that ε_i is orthogonal to $\mathbf{X}_{i_{st}}$ and beliefs. It is also assumed that ε_i adopts a logistic distribution.

$$\Lambda(\epsilon) \equiv \frac{e^{\epsilon}}{1 + e^{\epsilon}} \tag{2.18}$$

Beliefs Estimation

Let $\overline{\boldsymbol{\pi}}_{i_{st}} \equiv \{\overline{\pi}^S_{A,i_{st}}, \overline{\pi}^M_{A,i_{st}}, \overline{\pi}^L_{A,i_{st}}\}$, vector of beliefs; $\overline{\pi}^S_{A,i_{st}}$ means the probability that the *ith* small firm, which belongs to segment s at time t, will be aggressive. Parallel interpretation can be made to the other parameters. As we have seen, there are 3 periods, 9 SIC industries, and 3 types, then we have 81 probabilities which were distributed according to the i_{st}-th firm. So, in order to estimate this vector I follow the next strategy: first, the probability to be aggressive regardless types was calculated: $\overline{\boldsymbol{\pi}}_{A,i_{st}}$. Once this done, probabilities for each type were calculated as follow:

$$\overline{\boldsymbol{\pi}}_{i_{st}} = \begin{cases} \overline{\pi}^{S}_{A,i_{st}} = & \overline{\boldsymbol{\pi}}_{A,i_{st}} \times 1\{i_{st} \in Small \quad \} \\ \\ \overline{\pi}^{M}_{A,i_{st}} = & \overline{\boldsymbol{\pi}}_{A,i_{st}} \times 1\{i_{st} \in Medium\} \\ \\ \overline{\pi}^{L}_{A,i_{st}} = & \overline{\boldsymbol{\pi}}_{A,i_{st}} \end{cases}$$

Beliefs modeled this way capture behavior of small, medium and large firms, i.e., probability to be aggressive of each type of firm.

Conditional likelihood

Under the distributional assumption of ε, and using (2.8), conditional likelihood function (2.13) can be expressed as follows:

$$\begin{aligned} log\ f(y_{i_{st}}|\boldsymbol{X}_{i_{st}}, \overline{\boldsymbol{\pi}}_{i_{st}}, \boldsymbol{\theta}) \ &= \ \mathbf{1}\{y_{i_{st}} = 1\} log[\Lambda(\zeta_1 - \delta(\boldsymbol{X}_{i_{st}}, \boldsymbol{\theta}))] + \\ &+ \ \mathbf{1}\{y_{i_{st}} = 2\} log[\Lambda(\zeta_2 - \delta(\boldsymbol{X}_{i_{st}}, \boldsymbol{\theta})) - \Lambda(\zeta_1 - \delta(\boldsymbol{X}_{i_{st}}, \boldsymbol{\theta}))] + \\ &+ \ \mathbf{1}\{y_{i_{st}} = 3\} log[1 - \Lambda(\zeta_2 - \delta(\boldsymbol{X}_{i_{st}}, \boldsymbol{\theta}))] \end{aligned} \tag{2.19}$$

Then we will get the estimator by solving (for details, see the appendix):

$$\max_{\boldsymbol{\theta}, \{p_{i_{st}}\}_{i_{st}=1}^{N}} \ \sum_{i_{st}=1}^{N} log\ f(y_{i_{st}}|\boldsymbol{X}_{i_{st}}, \overline{\boldsymbol{\pi}}_{i_{st}}, \boldsymbol{\theta}) + \sum_{i_{st}=1}^{N} \ log\ p_{i_{st}} \tag{2.20}$$

subject to

$$p_{i_{st}} \geq 0, \ \sum_{i_{st}=1}^{N} p_{i_{st}} = 1, \ \sum_{i_{st}=1}^{N} p_{i_{st}} \Psi(\boldsymbol{X}_{i_{st}}, \theta) = 0 \tag{2.21}$$

Using Lagrange multipliers technique, is straightforward to show that

$$\sum_{i_{st}=1}^{N} log\ f(y_{i_{st}}|\boldsymbol{X}_{i_{st}}, \overline{\boldsymbol{\pi}}_{i_{st}}, \boldsymbol{\theta}) - \sum_{i_{st}=1}^{N} log(1 + \boldsymbol{\nu}' \Psi(\boldsymbol{X}_{i_{st}}, \boldsymbol{\theta})) - NlogN \tag{2.22}$$

where $\boldsymbol{\nu} \in \mathbb{R}^{81}$, are Lagrange multipliers.

Estimation results

NPV theory of capital investment predicts a positive coefficient for ΔQ_{ist}, however, economic theory does not provide a clear prediction for the sign of any remaining covariates in $\boldsymbol{X}_{ist}$. Extremum values of the sample were eliminated because they were source of bias.

By solving (2.22) we found next results summarized in Table 1 and Table 2. All Lagrange multipliers were equal to zero statistically significant, which means that model fits well:

Table 1. Estimation Results
for strategic coefficients
(Standard Errors in parentheses)

α^S	0.7989*
	(0.3574)
α^M	0.6356*
	(0.3082)
α^L	0.3122
	(0.3015)

(∗) Statistically significant at a 5% level.

The estimates for α^S and α^M were significant at 5% confidence level. Both were positive, which means that small and medium firms care about actions of their own type. For example, small firms will be aggressive if they believe that other small firms would be aggressive. Parallel analysis could be made for the medium size firms. On the other hand, the coefficient α^L was not statistically significant. It means that large firms are not affected by decisions made by other large firms.
The hypothesis test[9] $H_0 : \alpha^S + \alpha^L = 0$ was rejected, which means that small firms care about the large firms decisions tending to be aggressive if they believe that large

[9]All hypothesis test were made at 5% level of significance.

firms will be aggressive. The hypothesis of $H_0 : \alpha^M + \alpha^L = 0$ was rejected too. The analysis for medium firm is the same as in the small firm case. It can be notice that $|\alpha^S + \alpha^L| > |\alpha^M + \alpha^L|$, and $|\alpha^S| > |\alpha^M| > |\alpha^L|$. This means that small firms are more worried about others' decisions than medium and large size firms. Small firms would be aggressive if they believe that medium or large firms would be aggressive. That can be reasonable explained using "survivor" analysis.

Table 2. Estimation Results
for private information variables

(Standard Errors in parentheses)

$y_{i_{st-1}}$	-0.6310*
	(0.0774)
$\Delta\%S_{i_{st}}$	-0.1767*
	(0.0684)
$\Delta Q_{i_{st}}$	0.0937*
	(0.0240)
ζ_1	-1.5137*
	(0.2765)
ζ_2	-0.6005*
	(0.2721)

($*$) Statistically significant at a 5% level.

About private information variables, we can say that the sign of the coefficient $\Delta Q_{i_{st}}$ was significant at 5% confidence level and positive, as predicted by the NPV theory of investment. Coefficient of $\Delta\%S_{i_{st}}$, i.e., the variable that captures the short run behavior of the firms, has negative sign and is significant. It means that, at least in the short run, firms tend to be non aggressive, ceteris paribus. Finally, last period firms' behavior, $y_{i_{st-1}}$, has a negative significant coefficient. It means that past behavior of the firms, drive them to be non aggressive in the next period. It could be understood as an adjustment that firms make considering how they acted in the past, as if they correct in a conservative way using their past experiences. Finally, both cutoffs were negative and significant at 5% confidence level. It was rejected the hypothesis that $H_0 : \zeta_1 = \zeta_2$, which means that, just as it was said, there are three decisions to be made: passive, neutral and aggressive[10]. Given (2.5) and (2.6), we can conclude that firms start to make their decisions at a certain desutility level, specially those that decide to be passive or neutral. This confirms that being aggressive is the "best" state in this game.

2.5 Conclusions

Asymmetric information is the appropriate setting for a number of interaction 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 entrails 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 semi-empirical likelihood methods[11]. This allows us to simultaneously estimate payoff parameters, beliefs and 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

[10] If $\zeta_1 = \zeta_2$, it means that the model should be dichotomic, in other words, decisions variable only would take two actions: be passive or aggressive.

[11] Empirical means "non parametric".

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 and uniqueness of equilibrium are very important, and fortunately more tractable than they are in general in perfect information models.

An application for investment model was analyzed and estimated here. In this model we have three actions: firms decide to be passive, neutral or aggressive in the investment sense; and there are three types: small, medium and large firms. We analyzed how they interact with each other. We found evidence that small firms care the most about the other firms' action of their own and different type (medium and large). Same result can be applied to medium size firms but less strongly than small size. Large firms do not care about others' actions, maybe because they are "strong" and have certain "self-confidence" at the moment of make their investment decisions, or maybe the market structure helps them.

An extension of this model could be to deal with dynamic models in which the change of interaction coefficients ("$\alpha' s$") over time was permitted . In particular, this model can be extended to deal with panel data structure, but ordered response models in panel data are relatively difficult to estimate, because of the nonlinear structure of the model in which fixed effects do not disappear simply applying the first difference technique (Bo Honore (2002)).

Appendix C2.1

Existence of equilibria

To prove the existence of a solution of (2.5), note that the equilibrium conditions

$$\int_{\boldsymbol{x}} \Psi(\boldsymbol{\theta}, X) dG_{\boldsymbol{X}}(\boldsymbol{x}) = \mathbf{0}$$

can be expressed as

$$\begin{aligned}
\bar{\pi}_A^S &= \int_{\boldsymbol{x}} \frac{\sum_{i=1}^N [1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))] \mathbf{1}\{k = S\}}{\sum_{i=1}^N \mathbf{1}\{k = S\}} dG_{\boldsymbol{X}}(\boldsymbol{x}) \\
\bar{\pi}_A^M &= \int_{\boldsymbol{x}} \frac{\sum_{i=1}^N [1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))] \mathbf{1}\{k = M\}}{\sum_{i=1}^N \mathbf{1}\{k = M\}} dG_{\boldsymbol{X}}(\boldsymbol{x}) \\
\bar{\pi}_A^L &= \int_{\boldsymbol{x}} \frac{\sum_{i=1}^N [1 - F(\zeta_2 - \delta(\boldsymbol{\theta}, \boldsymbol{X}))] \mathbf{1}\{k = L\}}{\sum_{i=1}^N \mathbf{1}\{k = L\}} dG_{\boldsymbol{X}}(\boldsymbol{x})
\end{aligned}$$

Now assuming that the marginal distribution of ε is continuous, so the resulting probabilities are continuous (logistic distribution assumed here satisfy this condition). Then, for an arbitrary value of the parameter $\boldsymbol{\theta}_2$ the right hand side of the equation presented above is a continuous function of the left hand side vector, $\boldsymbol{\theta}_1$. Therefore, the right hand side is a continuous mapping form $[0,1]^3 \times [0,1]^3$ and by Brower's Fixed Point Theorem, it has a fixed point. Since this true for an arbitrary value of $\boldsymbol{\theta}_2$, it must hold for $\boldsymbol{\theta}_2^0$, the true vales of the parameters. This proves that an equilibrium exists.[12]

$\square$

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

Suppose the following conditions are satisfied.

Ap1.- All equilibrium beliefs are strictly between 0 and 1.

[12] In our application, it holds but the mapping is $[0,1]^{81} \times [0,1]^{81}$

Ap2.- Identification conditions discussed above are satisfied.

Ap3.- The log-likelihood log $f(Y|\boldsymbol{X},\boldsymbol{\theta})$ satisfy the usual technical conditions for asymptotic consistency and normality of MLE.

Ap4.- The sample jacobian matrix for equilibrium conditions $\frac{1}{N}\sum_{i=1}^{N}\nabla_{\boldsymbol{\theta}}\Psi(\boldsymbol{x}_i,\boldsymbol{\theta})$, converges uniformly in probability to its expected value if $\boldsymbol{\theta}$ converges to $\boldsymbol{\theta}_0$.

A5.- Technical conditions for the asymptotic normality of $\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\nabla_{\boldsymbol{\theta}}\Psi(\boldsymbol{x}_i,\boldsymbol{\theta}_0)$ are satisfy.

Let

$$\mathfrak{I}_0 = Var[f(Y|\boldsymbol{X},\boldsymbol{\theta}_0)],\ A_0 = E[\nabla_{\boldsymbol{\theta}}\Psi(\boldsymbol{x}_i,\boldsymbol{\theta}_0)],\ B_0 = E[\Psi(\boldsymbol{x}_i,\boldsymbol{\theta}_0)\Psi(\boldsymbol{x}_i,\boldsymbol{\theta}_0)']$$

Then we have that:

$$\sqrt{N}(\hat{\boldsymbol{\theta}}^{EL} - \boldsymbol{\theta}_0) \to^d N(\mathbf{0},\boldsymbol{\Omega})$$

Where

$$\boldsymbol{\Omega} = (\mathfrak{I}_0 + A_0' B_0^{-1} A_0)^{-1}$$

Proof:

The corresponding Lagrangian for the EL estimation problem is given by

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{N} log\ f(y_i|\boldsymbol{x}_i,\boldsymbol{\theta}) + \sum_{i=1}^{N} p_i + \\ &+ \lambda(1-\sum_{i=1}^{N} p_i) - N\boldsymbol{\nu}'\sum_{i=1}^{N} p_i\Psi(\boldsymbol{x}_i,\boldsymbol{\theta})\end{aligned}$$

$\lambda \in \mathbb{R}$ *and* $\boldsymbol{\nu} \in \mathbb{R}^3$ are lagrange multipliers.

F.O.C. with respect to p_i yield

$\lambda = N$ and $p_i = \frac{1}{N(1+\boldsymbol{\nu}'\Psi(\boldsymbol{x}_i|\boldsymbol{\theta}))}$, $i = \{1, 2, ..., N\}$.

Plug-in back p_i in our moment condition $\sum_{i=1}^{N} p_i \Psi(\boldsymbol{x}_i, \boldsymbol{\theta}) = 0$, we have:

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\Psi(\boldsymbol{x}_i, \boldsymbol{\theta})}{1+\boldsymbol{\nu}'\Psi(\boldsymbol{x}_i, \boldsymbol{\theta})} = 0$$

Solving this non linear equation, we can determine: $\hat{\boldsymbol{\nu}}(\boldsymbol{\theta})$. Equivalently, this Lagrange multipliers can be found solving the next minimization problem:

$$\min_{\boldsymbol{\nu}\in\mathbb{R}^3} - \sum_{i_{st}=1}^{N} \log(1+\boldsymbol{\nu}'\Psi(\boldsymbol{x}_i, \boldsymbol{\theta}))$$

Then, we can obtain $\hat{p}_i$, and plugging back in to the joint semi-empirical likelihood:

$$\sum_{i_{st}=1}^{N} \log f(y_i|\boldsymbol{x}_i, \boldsymbol{\theta}) - \sum_{i=1}^{N} \log(1+\boldsymbol{\nu}'\Psi(\boldsymbol{x}_i, \boldsymbol{\theta})) - N \log N$$

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

$$\begin{aligned} S_{1,N}(\hat{\boldsymbol{\theta}}^{EL}, \boldsymbol{\nu}) &\equiv \sum_{i=1}^{N} \nabla_\theta \log f(y_i|\boldsymbol{x}_i, \hat{\boldsymbol{\theta}}^{EL}) - \frac{1}{N}\sum_{i=1}^{N}\frac{\nabla_\theta \Psi(\boldsymbol{x}_i, \hat{\boldsymbol{\theta}}^{EL})'\boldsymbol{\nu}}{1+\boldsymbol{\nu}'\Psi(\boldsymbol{x}_i, \hat{\boldsymbol{\theta}}^{EL})} = 0 \\ S_{2,N}(\hat{\boldsymbol{\theta}}^{EL}, \boldsymbol{\nu}) &\equiv \frac{1}{N}\sum_{i=1}^{N}\frac{\Psi(\boldsymbol{x}_i, \hat{\boldsymbol{\theta}}^{EL})}{1+\boldsymbol{\nu}'\Psi(\boldsymbol{x}_i, \hat{\boldsymbol{\theta}}^{EL})} = 0 \end{aligned}$$

Which means that solving (2.23), we can obtain the expected estimators. Using this estimators, we can determine the asymptotic properties of them as follows.

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

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

Where,

$$\begin{aligned} S_{1,N}^0 &= \frac{1}{N}\sum_{i=1}^{N} \nabla_{\boldsymbol{\theta}} \log f(y_i|\boldsymbol{x}_i, \boldsymbol{\theta}_0) \\ S_{2,N}^0 &= \frac{1}{N}\sum_{i=1}^{N} \Psi(\boldsymbol{x}_i, \boldsymbol{\theta}_0) \\ I_N &= \frac{1}{N}\sum_{i=1}^{N} \nabla_{\boldsymbol{\theta},\boldsymbol{\theta}'} \log f(y_i|\boldsymbol{x}_i, \boldsymbol{\theta}_0) \\ A_N &= \frac{1}{N}\sum_{i=1}^{N} \nabla_{\theta} \Psi(\boldsymbol{x}_i, \boldsymbol{\theta}_0) \\ B_N &= \frac{1}{N}\sum_{i=1}^{N} \Psi(\boldsymbol{x}_i, \boldsymbol{\theta}_0)\Psi(\boldsymbol{x}_i, \boldsymbol{\theta}_0)' \end{aligned}$$

Then, under regularity conditions, we have:

$$I_N \rightarrow^p \Im_0, \quad A_N \rightarrow^p A_0, \quad B_N \rightarrow^p B_0$$

and

$$\begin{matrix} \sqrt{N} S_{1,N}^0 \\ \sqrt{N} S_{2,N}^0 \end{matrix} \rightarrow^d \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \ where, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \Im_0 & 0 \\ 0 & B_0 \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} \sqrt{N}(\hat{\boldsymbol{\theta}}^{EL} - \boldsymbol{\theta}_0) \\ \sqrt{N}\boldsymbol{\nu} \end{pmatrix} \rightarrow^d \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$$

Where,

$$\boldsymbol{\Omega} = \begin{pmatrix} -\Im_0 & -A_0' \\ A_0 & -B_0 \end{pmatrix}^{-1} \begin{pmatrix} \Im_0 & 0 \\ 0 & B_0 \end{pmatrix} \begin{pmatrix} -\Im_0 & -A_0' \\ A_0 & -B_0 \end{pmatrix}^{-1'}$$

and so we get

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

As we claimed. $\square$

Appendix C2.2

Empirical Strategy

In order to reach convergency in the minimization problem exposed above, we can use the logarithm function proposed by Owen (2001):

$$log_*(z) = \begin{cases} log(z) & if \ z > \varepsilon \\ \\ log(\varepsilon) - 1.5 + z/\varepsilon - z^2/(z\varepsilon^2) & if \ z \leq \varepsilon \end{cases}$$

for some $\varepsilon > 0$ (one recommended to use is $\varepsilon = \frac{1}{N}$)

At the same time, the initial values of $\boldsymbol{\nu}$ were values near to 0. Results are presented at the next page.

SIC Sector Description

SIC	N_s	%
2834	222	22.93
2836	86	8.88
3674	138	14.26
3845	146	15.08
2911	83	8.57
3312	80	8.26
3559	67	6.92
3089	54	5.58
3714	92	9.5
Total	968	

2834: Tech Segment 1. Pharmaceutical Preparations.

2836: Tech Segment 1. Biological Products, Except Diagnostic Substances.

3674: Tech Segment 1. Semiconductors and Related Devices.

3845: Tech Segment 1. Electromedical and Electrotherapeutic Apparatus.

2911: Tech Segment 2. Petroleum Refining.

3312: Tech Segment 2. Steel Works, Blast Furnaces (Including Coke Ovens), and Rolling Mills.

3559: Tech Segment 2. Special Industry Machinery, Not Elsewhere Classified.

3089: Tech Segment 3. Plastics Products, Not Elsewhere Classified.

3714: Tech Segment 3. Motor Vehicle Parts and Accessories.

Source of Variables (from COMPUSTAT North America)

data4: Current Assets Total.

data8: Property, Plant, and Equipment-Total (Net).

data12: Sales (Net) Total.

data24: Price-Close.

data25: Common Shares Outstanding.

data29: Employees.

data30: Property, Plant, and Equipment-Capital Expenditure (Schedule V).

data33: Intangibles.

data34: Debt in Current Liabilities.

data60: Common Equity-Total.

data130: Preferred Stock-Carrying Value.

Chapter 3

Technical Efficiency in the Mexican Manufacturing Sector: A Stochastic Frontier Approach

3.1 Introduction

A basic assumption in microeconomic theory is that firms, in general, are homogeneous. Such is represented by the perfect competition framework in which all the firms are assumed to operate at the same level of efficiency. Nonetheless, there are many studies that have shown precisely the contrary (Caves 1989), and Mexico is not the exception to such finding. Moreover, in the last times firms have gradually been exposed to a strong open economy in the worldwide and Mexican firms too. One could expect an improvement in the development of firms which were exposed into the competition environment. In fact, NAFTA (North American Free Trade Agreement), had this as one of its main purposes. But, in practice, what happened? This paper analysis the technical efficiency in the Mexican manufacturing sector in which determinants and changes of the efficiency from the beginning of NAFTA, 1994 to 2001, are measured using the Panel Data Stochastic Frontier Models.

NAFTA transformed Mexico from an inward-looking economy into largely open economy (Calderon and Voicu, 2004). Tariff politics were designed in order to open Manufacturing sector gradually giving it the opportunity to have more competitive firms. A good indicator that could help to measure that impact is the productivity (efficiency) at firm level and its evolution. There are many ways for modeling it. Calderon and Voicu (2004), for example, studied a detailed analysis of performance of Mexican manufacturing firms between 1993 and 2000. They constructed the estimators of individual plant productivity and investigated the relationship between trade reforms and plant performance, using Levinshon and Petrin (2001) methodology instead of Olley and Pakes (1996) which seems to be more restrictive. They found that “access to imported inputs is more significant vehicle for productivity enhancing effects of trade openness, and that investment in technology is, by far, most strongly correlated with plant productivity”.

In this work we used the main information gathered in Annual Industrial Survey (AIS) which allows us modeling efficiency (productivity) using Panel Data Stochastic Frontier technique developed below, in order to show determinants of the poor development observed by Mexican manufacturing sector, in contrast to the optimistic projections that were made. At the same time, we try to explain that the crisis was not the main reason of this poor development, but the lack or lost of efficiency observed. A good survey of studies realized using the AIS for Mexico, can be found in Calderon and Voicu (2004). However, none of them has used the methodology presented here.

3.2 Panel Data Stochastic Frontier Model

The field of the stochastic frontier estimation of technical (and cost) efficiency is enormous and growing (Greene 2002). Most of studies are based on the fixed effects model (Schmidt and Sickles 1984) and random effects model (Pitt and Lee 1981). In both cases time invariant technical efficiency is assumed. This could be questionable, particularly in a long panel. As an alternative approach, the Battese and Coelli's (1992) parametrization of time-effects has been proposed. Following (Kumbahakar and Lovell, 2000, pg. 63-114), I present a brief review of stochastic frontier models in order to estimate the technical efficiency.

3.2.1 Cross-Sectional Production Frontier Models

It is assumed, in general that cross-sectional data on the quantities of N inputs used to produce a *single output* are available to the econometrist for each of I producers. A production frontier model can be written as

$$y_i = f(x_i; \beta) \cdot TE_i \tag{3.1}$$

where y_i is the scalar product output of the producer i, $i = 1, ..., I$; x_i is a vector of N inputs used by producer i; β is a vector of technology parameters to be estimated, and $f(x_i; \beta)$ is the production frontier; in other words, $f(\cdot)$ measure the possibility to reach the maximum product given different combinations of inputs and technological parameters, β.

Then, rearranging (3.1), we have

$$TE_i = \frac{y_i}{f(x_i; \beta)} \tag{3.2}$$

which defines technical efficiency as the ratio of observed output to maximum feasible output. Indeed, y_i achieves its maximum feasible value of $f(x_i; \beta)$ if, and only if, $TE_i = 1$. Otherwise $TE_i < 1$ provides a measure of the shortfall of observed output from maximum feasible. It is important to notice that $f(x_i; \beta)$ in (3.1) is *deterministic*, which means that if there is a shortfall in (3.2) should be attributed, directly, to the inefficiency. This model does not capture some random shocks, that could explain that shortfall in the production process, shocks that are not under producers' control. To incorporate producer-specific random shocks into the analysis requires the specification of a *stochastic production frontier* as follows

$$y_i = f(x_i;\beta) \cdot exp\{v_i\} \cdot TE_i \tag{3.3}$$

where $[f(x_i;\beta) \cdot exp\{v_i\}]$ is the *stochastic production frontier*, which is defined for two parts: the *deterministic* one, $f(x_i;\beta)$; and, the *stochastic*[1] one: $exp\{v_i\}$. Then, *technical efficiency* can be represented in this way

$$TE_i = \frac{y_i}{f(x_i;\beta) \cdot exp\{v_i\}} \tag{3.4}$$

here, y_i achieves its maximum feasible output of $[f(x_i;\beta) \cdot exp\{v_i\}]$ if, and only if, $TE_i = 1$. Otherwise $TE_i < 1$ provides a measure of the shortfall of observed output from maximum feasible output in an environment characterized by $exp\{v_i\}$. Technical efficiency can be estimated using either the deterministic production frontier model given by equations (3.1) and (3.2), or the stochastic production frontier model given by the equations (3.3) and (3.4).

The goal is to estimate the technical parameters, $\beta' s$, and the technical efficient measure, TE_i. There is more than one way to achieve this objective. Here, I will mention some estimation techniques in the cross-section case, in order to be deeper in panel data stochastic frontier analysis case, which will be cover with more detail. Then, cross-sectional frontier model can be estimated as follows[2]:

$$Deterministic\ Production\ Frontier \begin{cases} 1 & \text{Goal Programing} \\ \\ 2 & \text{Corrected Ordinary Least Squares} \\ \\ 3 & \text{Modified Ordinary Least Square} \end{cases}$$

[1]Note: assumptions about this component will be specified below.

[2]For details, see Kumbhakar and Lovell, 2000. Pg. 66-95.

$$\textit{Stochastic Production Frontier}\begin{cases} 1 \quad \text{Normal-Half Normal Model} \\ \\ 2 \quad \text{Normal-Exponential Model} \\ \\ 3 \quad \text{Normal-Gamma Model} \\ \\ 4 \quad \text{Method of Moments Approach} \end{cases}$$

3.2.2 Panel Data Production Frontier Models

Evidently, panel data (repeated observation on each producer, or, the same producer followed in more than one period) contains more information than does a single cross section. Have access to panel data is convenient in more than one sense: First, conventional panel data techniques can be adapted in order to estimate stochastic production frontier models. Second, repeated observations on a sample of producer can serve as a substitute for strong assumptions made in the cross-sectional environment. Finally, since adding more observations on each producer generates information not provided by adding more producers to a cross-section, the technical efficiency of each producer in the sample can be estimated consistently as $T \to \infty$, T being the number of observations on each producer.

Panel data can be *balanced* (each producer is observed T times) and *unbalanced* (producer i is observed $T_i \leq T$). In this study we use a balanced panel. Again, we assume that there is *more than one inputs* (multiple inputs) that are combined using certain technology (represented by the production function), which result is a *single output*. There are *two main assumptions* to be done having in hand a panel data: Can be allowed that technical efficiency vary across producer, but is assumed to be constant through time for each producer, this model is known as *time invariant technical efficiency*; this assumption could be implausible in long panels. That is why we include the assumption of *time variant technical efficiency* which allows that technical efficiency vary across producer and through time for each producer.

Time-Invariant Technical Efficiency

Set up of the model: we assumed to have I producers ($i = 1, ..., I$), followed in T periods ($t = 1, ...T$). A Cobb-Douglas production frontier with *time-invariant* technical efficiency:

$$lny_{it} = \beta_0 + \sum_n \beta_n lnx_{nit} + v_{it} - u_i \tag{3.5}$$

where:

y_{it}: is the output of the producer i at time t,

β_0: is the intercept,

β_n: are the "technological parameters",

x_{nit}: is the vector of inputs of the production function of producer i at time t,

v_{it}: two-sided "noise" component. Production can be affected by random shocks out side the control of producers.

u_i: shocks attributed to the technical efficiency.

Then, v_{it} represents random statistical noise and $u_i \geq 0$ represents technical inefficiency[3]. Notice that technical change is not allowed, since u_i does not vary over the time, but vary over producers. This model is very similar to a conventional panel data model with producer effects but without time effects, the only difference is that producer effects are required to be nonnegative. Again, parameters of the model, and technical efficiency can be estimated in a number of ways.

[3]Notices that: $lny_{it} - (\beta_0 + \sum_n \beta_n lnx_{nit} + v_{it}) = -u_i$, but $lny_{it} \leq (\beta_0 + \sum_n \beta_n lnx_{nit} + v_{it})$, which implies that u_i must be positive.

The Fixed-Effects Model: Assumptions:

1. $u_i \geq 0$
2. v_{it} are iid $(0, \sigma_v^2)$
3. We make no distributional assumptions on the u_i
4. We allowed u_i to be correlated with regressors or with the v_{it}.

Given that u_i does not vary in time, it is treated as fixed (nonrandom) effects, then, can be considered as *specific intercept parameters*, which can be estimated along with the $\beta_n s$. Consequently, the model can be estimated by applying OLS to:

$$lny_{it} = \beta_{0i} + \sum_n \beta_n lnx_{nit} + v_{it} \tag{3.6}$$

where $\beta_{0i} = \beta_0 - u_i$ are producer specific intercepts. After the estimation we can employ the normalization

$$\hat{\beta}_0 = max_i\{\hat{\beta}_{0i}\} \tag{3.7}$$

then, u_i are estimated using

$$\hat{u}_i = \hat{\beta}_0 - \hat{\beta}_{0i} \tag{3.8}$$

notice that this ensures the assumption that $u_i \geq 0$. Producer-specific estimates of technical efficiency are then given by

$$TE_i = exp\{-\hat{u}_i\} \tag{3.9}$$

we can observe that in this model at least one producer is assumed to be 100% technical efficient, and the rest of producers measure their efficiency relatively to this "*benchmark*" producer(s).

Fixed-effect model is quite simple to be calculated, and has nice consistency properties, and provides consistent estimates of producer-specific technical efficiency.

Nonetheless, fixed-effects model have some potentially drawback: u_i not necessarily capture only the time-invariant technical efficiency, capture *all* phenomena (such as the regulatory environment, as an example). Then, the econometrist can confound variation with technical efficiency with variation in other effects. That is why in the literature was proposed the next model.

The Random-Effects Model: In this framework we assumed that u_i is randomly distributed with constant media and variance, but uncorrelated with the regressors and the error term v_{it}. There is no another assumption to be done about u_i, only that it should hold the nonnegative requirement. Again, it is assumed that v_{it} have zero expectation and constant variance. Under this assumptions we are able to include *time-invariant* regressors in the model.

$$\begin{aligned} \ln y_{it} &= [\beta_0 - E(u_i)] + \sum_n \beta_n \ln x_{nit} + v_{it} - [u_i - E(u_i)] \\ &= \beta_0^* + \sum_n \beta_n \ln x_{nit} + v_{it} - u_i^* \end{aligned} \tag{3.10}$$

This random-effects model fits exactly into the one-way error components model in the panel data literature (see Baltagi 2005 pg. 107-8), then can be estimated by the standard two-step Generalized Least Squares (GLS) method. In the first step all parameters are estimated using OLS. The two variance components are estimated by any of several methods. In the second step β_0^* and the $\beta_n s$ are reestimated using feasible GLS. There is only one intercept term to be estimated because β_0^* does not depend on i, because by assumption $E(u_i)$ is a constant. Once β_0^* and $\beta_n s$ have been estimated using feasible GLS, the u_i^* can be estimated from the residuals by means of

$$\hat{u}_i^* = \frac{1}{T} \sum_t \left[\ln y_{it} - \hat{\beta}_0^* - \sum_n \hat{\beta}_n \ln x_{nit} \right] \tag{3.11}$$

And finally, the estimations of u_i are obtained by means of the normalization:

$$\hat{u}_i = max_i\{\hat{u}_i^*\} - \hat{u}_i^* \tag{3.12}$$

These estimates are consistent as both $I \rightarrow \infty$ and $T \rightarrow \infty$. Estimates of producer-specific technical efficiency can be obtained by substituting $\hat{u}_i$ in (3.9). There is more than one ways of estimates u_i, for example, using best linear unbiased predictor (BLUP), (see Kumbahakar and Lovell, 2000, pg. 101).

Finally, in the time-invariant framework we can assume certain distributions of the errors and estimate parameters and technical efficiency using *maximum likelihood*.

Maximum Likelihood: This technique is widely used in empirical analysis. In this work was used. The general setup of the model is:

(i) $v_{it} \sim iid\ N(0, \sigma_v^2)$

(ii) $u_i \sim iid\ N^+(0, \sigma_u^2)$

(iii) v_{it} and u_i are distributed independently of each other, and of the regressors.

Pit and Lee (1981) used this assumptions to estimate technical efficiency using panel data. We will use this parametric specification of the *random effects* model which adds the normality and half-normality assumptions[4], considering the inefficiency as *time-invariant*. The density of u is given by

$$f(u) = \frac{2}{\sqrt{2\pi}\sigma_u} exp\Big\{ -\frac{u^2}{2\sigma_u^2} \Big\} \tag{3.13}$$

the density function of $\mathbf{v} = (v_1, ..., v_T)'$, which depends on time, is

$$f(\mathbf{v}) = \frac{1}{(2\pi)^{T/2}\sigma_v^T} \cdot exp\Big\{ \frac{-\mathbf{v'v}}{2\sigma_v^2} \Big\} \tag{3.14}$$

then, given the independence assumption the joint density function of u and $\mathbf{v}$ is

[4] N^+ represents the normal distribution when the support of the error term u_i is positive.

$$f(u, \mathbf{v}) = \frac{2}{(2\pi)^{(T+1)/2}\sigma_u\sigma_v^T} \cdot exp\Big\{ - \frac{u^2}{2\sigma_u^2} - \frac{\mathbf{v'v}}{2\sigma_v^2}\Big\} \tag{3.15}$$

the joint density of u and $\boldsymbol{\varepsilon} = (v_1 - u, ..., v_T - u)'$ is

$$f(u, \boldsymbol{\varepsilon}) = \frac{2}{(2\pi)^{(T+1)/2}\sigma_u\sigma_v^T} \cdot exp\Big\{ - \frac{(u-\mu_*)^2}{2\sigma_*^2} - \frac{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{2\sigma_v^2} + \frac{\mu_*^2}{2\sigma_*^2}\Big\} \tag{3.16}$$

where

$$\mu_* = -\frac{T\sigma_u^2\overline{\boldsymbol{\varepsilon}}}{\sigma_v^2 + T\sigma_u^2}$$

$$\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma_v^2 + T\sigma_u^2}$$

$$\overline{\boldsymbol{\varepsilon}} = \frac{1}{T}\sum_t \boldsymbol{\varepsilon}_{it}$$

consequently, the marginal density function of $\boldsymbol{\varepsilon}$ is

$$\begin{aligned} f(\boldsymbol{\varepsilon}) &= \int_0^\infty f(u, \boldsymbol{\varepsilon}) du && (3.17) \\ &= \frac{2[1 - \boldsymbol{\Phi}(-\mu_*/\sigma_*)]}{(2\pi)^{T/2}\sigma_v^{(T-1)}(\sigma_v^2 + T\sigma_u^2)^{1/2}} \cdot exp\Big\{ - \frac{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{2\sigma_v^2} + \frac{\mu_*^2}{2\sigma_*^2}\Big\} && (3.18) \end{aligned}$$

where $\boldsymbol{\Phi}(\cdot)$ is the standard normal cumulative distribution. Then, we assumed that the econometrist have in hand sample of I producer, each observed at for T periods of time, so the likelihood function is

$$\begin{aligned} \ln L &= constant - \frac{I(T-1)}{2}\ln\sigma_v^2 - \frac{I}{2}\ln(\sigma_v^2 + T\sigma_u^2) \\ &+ \sum_i \ln\Big[1 - \boldsymbol{\Phi}\Big(- \frac{\mu_{*i}}{\sigma_*}\Big)\Big] - \frac{\sum_i \boldsymbol{\varepsilon}_i'\boldsymbol{\varepsilon}_i}{2\sigma_v^2} + \frac{1}{2}\sum_i \Big(\frac{\mu_{*i}}{\sigma_*}\Big)^2 && (3.19) \end{aligned}$$

This log likelihood function can be maximized with respect to the parameters to obtain maximum likelihood[5] estimates of β, σ_v^2 and σ_u^2. Next step is to obtain estimates of *producer-specific time-invariant technical efficiency.* We start deriving the conditional distribution $(u|\boldsymbol{\varepsilon})$, using its definition:

$$\begin{aligned} f(u|\boldsymbol{\varepsilon}) &= \frac{f(u,\boldsymbol{\varepsilon})}{f(\boldsymbol{\varepsilon})} \\ &= \frac{1}{(2\pi)^{1/2}\sigma_*[1-\boldsymbol{\Phi}(-\mu_*/\sigma_*)]} \cdot exp\left\{-\frac{(u-\mu_*)^2}{2\sigma_*^2}\right\} \end{aligned} \tag{3.20}$$

which is the density function of a variable distributed as $N^+(\mu_*, \sigma_*^2)$, where N^+ indicates that is positive normal distribution. Then, the mean (or the mode) of this distribution can be used as a point estimator of technical efficiency, then we have:

$$\hat{u}_i = E(u_i|\boldsymbol{\varepsilon}_i) = \mu_{*i} + \sigma_* \left[\frac{\boldsymbol{\phi}(-\mu_{*i}/\sigma_*)}{1-\boldsymbol{\Phi}(-\mu_{*i}/\sigma_*)}\right] \tag{3.21}$$

The estimators of u_i are consistent as $T \to \infty$. And, again, $\hat{u}_i$ can be substituted in the equation (3.9) in order to obtain the producer-specific estimates of time-invariant technical efficiency.

Time-Variant Technical Efficiency

If the econometrist have access to a long panel, it is plausible to think that technical efficiency is not constant. Particularly in a competitive environment. Then, we expect that technical inefficiency changes over time. Then, we are able to relax the assumptions that the producer-specific technical efficiency is *time-variant.* As in time-invariant model, the estimation of a time-varying technical efficiency model can be reach using fixed or random effects and maximum likelihood approach.

[5]Remember that we are in the "second step" in which we have already estimated u and v, then "observable" variables in this likelihood function are estimation errors that come from observable data (x and y), ε.

Fixed-Effects Models and Random-Effects Models: Cornwell, Schmidt, and Sickles (1990), and Kumbhakar (1990) were the first to propose a stochastic production frontier panel data model with time-varying technical efficiency.

$$\begin{aligned}\ln y_{it} &= \beta_{0t} + \sum_n \beta_n \ln x_{nit} + v_{it} - u_{it} \\ &= \beta_{it} + \sum_n \beta_n \ln x_{nit} + v_{it} \end{aligned} \tag{3.22}$$

where β_{0t} is the production frontier intercept common to all producers in period t, $\beta_{it} = \beta_{0t} - u_{it}$ is the intercept for producer i in period t, an all other variables are as previously defined. The objective is to obtain the estimates of the parameters describing the structure of production technology, and the second objective is to obtain producer-specific estimates of technical efficiency. The main problem is the identification of the intercept, in order to reduce the amount of $I \cdot T$ intercepts to another amount handle, Cornwell, Schmidt, and Sickles (1990) addressed this problem by specifying

$$\beta_{it} = \Omega_{i1} + \Omega_{i2}t + \Omega_{i3}t^2 \tag{3.23}$$

which reduces the number of intercept parameters to I*3. But, most importantly is that this specification allows technical efficiency to vary through time, and in a different manner for each producer. We can delete u_{it} from (3.22), estimate the $\beta_n s$, from the residuals, and regress the residuals on a constant t and t^2 to obtain estimations of $(\Omega_{i1}, \Omega_{i2}, \Omega_{i3})$ for each producer. Then we can estimate β_{it} and can be defined $\hat{\beta}_{0t} = max_i\{\hat{\beta}_{it}\}$ as the estimated intercept of the production frontier in the period t. The technical efficiency of each producer in period t is then estimated as $TE_{it} = exp\{-\hat{u}_{it}\}$, where $\hat{u}_{it} = (\hat{\beta}_{0t} - \hat{\beta}_{it})$. Thus, similar that in the time-invariant case, in each period at least one producer is consider to be 100% technical efficiency, but it can change through time.

Maximum Likelihood: This procedure is too similar to the time-invariant, then we arrive to the next likelihood function[6]:

$$
\begin{aligned}
\ln L &= constant - \frac{I}{2}\ln\sigma_*^2 - \frac{1}{2}\sum_i a_{*i} - \frac{I \cdot T}{2}\ln\sigma_v^2 \\
&- \frac{I}{2}\ln\sigma_u^2 + \sum_i \ln\left[1 - \mathbf{\Phi}\Big(-\frac{\mu_{*i}}{\sigma_*}\Big)\right]
\end{aligned} \tag{3.24}
$$

where:

$$\mu_{*i} = \frac{(\sum_t \beta_t \boldsymbol{\varepsilon}_{it})\sigma_v^2}{(\sigma_v^2 + \sigma_u^2 \sum_t \beta_t^2)}$$

$$\sigma_* = \frac{\sigma_v^2 \sigma_u^2}{\sigma_v^2 + \sigma_u^2 \sum_t \beta_t^2}$$

$$a_{*i} = \frac{1}{\sigma_v^2}\left[\sum_t \boldsymbol{\varepsilon}_{it}^2 - \frac{\sigma_u^2(\sum_t \beta_t \cdot \boldsymbol{\varepsilon}_{it})^2}{\sigma_v^2 + \sigma_u^2 \sum_t \beta_t^2}\right]$$

maximizing the log-likelihood function, ln L, we can estimate β, β_t, σ_v^2 and σ_u^2. Analogously, we can derive $u_i|\boldsymbol{\varepsilon}_i \sim N^+(\mu_{*i}, \sigma_*^2)$, and an estimator of u_i can be obtained from the mean (or the mode) of $u_i|\boldsymbol{\varepsilon}_i$.

$$\hat{u}_i = E(u_i|\boldsymbol{\varepsilon}_i) = \mu_{*i} + \sigma_*\left[\frac{\boldsymbol{\phi}(-\mu_{*i}/\sigma_*)}{1 - \mathbf{\Phi}(-\mu_{*i}/\sigma_*)}\right] \tag{3.25}$$

Finally, an alternative time-varying technical efficiency models was proposed by Battese and Coelli (1992), the model is based in next equations:

$$
\begin{aligned}
\ln y_{it} &= \beta_{0t} + \sum_n \beta_n \ln x_{nit} + v_{it} - u_{it} \\
&= \beta_{it} + \sum_n \beta_n \ln x_{nit} + v_{it}
\end{aligned} \tag{3.26}
$$

[6]For details see Kumbahakar and Lovell, 2000, pg. 110-113.

where

$$u_{it} = \beta(t) \cdot u_t \tag{3.27}$$

more than one author have proposed a particular functional[7] to $\beta(t)$, but in this work we followed the specification proposed by Battese and Coelli (1992):

$$\beta(t) = \exp\{-\gamma(t - T)\} \tag{3.28}$$

which has another parameter that should be estimated, γ. The function $\beta(t)$ satisfies the properties: (i) $\beta(t) \geq 0$ and $\beta(t)$ decreases at an increasing rate if $\gamma > 0$, and increases at a decreasing rate if $\gamma < 0$, or remains constant if $\gamma = 0$. Distributional assumptions are normal for v_{it} and truncated normal for u_i, and is used maximum likelihood to obtain estimates of all parameters in the model. The log-likelihood and its partial derivatives are in their paper, they showed that $(u_i|\boldsymbol{\varepsilon}_i) \sim iid\ N^+(\mu_{**i}, \sigma_*^2)$, where $\boldsymbol{\varepsilon}_i = \mathbf{v}_i - \beta \cdot u_i$ and

$$\mu_{**i} = \frac{\mu\sigma_v^2 - \beta'\boldsymbol{\varepsilon}_i\sigma_u^2}{\sigma_v^2 + \beta'\beta\sigma_u^2}$$

$$\sigma_*^2 = \frac{\sigma_u^2\sigma_v^2}{\sigma_v^2 + \beta'\beta\sigma_u^2}$$

$$\beta' = (\beta(1), ..., \beta(T))$$

if $\gamma = 0$ which implies that $\beta(t) = 1$, and $\beta'\beta = T$, technical efficiency is *time invariant* and μ_{**i} and σ_*^2 collapses to their time invariant version described in (3.16).

The minimum square error predictor of technical efficiency is

$$\begin{aligned} E(\exp\{-u_{it}\}|\boldsymbol{\varepsilon}_i) &= E(\exp\{\beta(t) \cdot u_t\}|\boldsymbol{\varepsilon}_i) \\ &= \frac{1 - \mathbf{\Phi}(\beta(t)\sigma_* - \mu_{*i}/\sigma_*)}{1 - \mathbf{\Phi}(-\mu_{*i}/\sigma_*)} \\ &\cdot \exp\Big\{ - \beta(t)\mu_{*i} + \frac{1}{2}\beta(t)^2\sigma_*^2\Big\} \end{aligned} \tag{3.29}$$

[7]Lee and Schmidt (1993).

In this paper we will compare both: Pit and Lee (1981), and Battese and Coelli's (1992) approaches. Results from two models will be compared. In the two specifications alike, we will use maximum likelihood estimator instead of least squares; the standard errors where corrected using bootstrap techniques[8], with 1000 replications.

3.2.3 The Model

The functional form in this work, for the sake of parsimony, will be Cobb-Douglas as was described in the last section[9]; despite its simplicity, it has proved to be a surprisingly good description of technology (Hayashi, 2000. Pg. 63). Then the model is:

Production Function

$$y_{it} = f(l_{it}, k_{it}, o_{it}) + v_{it} - u_{it} \tag{3.30}$$

$$i = 1, 2, ..., 4348;$$

$$t = 1, 2, ..., 8, (1994, ..., 2001)$$

Where[10],

$$f(l_{it}, k_{it}, o_{it}) = a + \beta_1 l_{it} + \beta_1 k_{it} + \beta_1 o_{it} \tag{3.31}$$

[8]For bootstrapping techniques, a useful guide is Handbook of Econometrics, 4, *Methodology and theory for the bootstrap* P. Hall (1994).

[9]The analysis could be done using a deterministic transcendental logarithmic (translog) production function (Greene 1997):

$$lny_{it} = ln\alpha_i + \lambda_t + \sum \beta_k ln_{kit} + \sum \beta_{2k}(lnx_{kit})^2 + \tfrac{1}{2}\Sigma_{q \neq w}\gamma_{qw}(lnx_{qit})(lnx_{wit}) + \varepsilon_{it}$$
$$where, k = 1, ..., p; i = 1, ..., N; t = 1, ..., t_i \text{ and } q = 1, ..., p; w = 1, ..., p, q \neq w$$

[10]$F(L, K, O) = AL^{\beta_1}K^{\beta_2}O^{\beta_3}$, *such that* $\beta_1 + \beta_2 + \beta_3 = 1;$ *then*, $lnF(\bullet) = f(l, k, o) = lnA + \beta_1 lnL + \beta_2 lnK + \beta_3 lnO = \alpha + \beta_1 l + \beta_2 k + \beta_3 o$, *where* $\alpha = lnA, l = lnL, k = lnK,$ *and* $o = lnO$

y is log of net value of total sales; l, log of value of the labor force; k, log of value of the net capital stock; o, log of value of other inputs, including value of the electrical consumption (see Section 3 below). It is important to recall that the error term u will be: u_i, in the *time invariant* model, and u_{it}, in the *time varying* one.

Inefficiency Analysis

In order to identify the sources of inefficiency, as a second step, u_{it} (estimated) is modeled using OLS. Explanatory variables are the following ones: 3 dummy variables and 2 decision variables.

$$\hat{u}_{it} = g(h_{it}) + W_{it} \tag{3.32}$$

$$h_{it} = (Export_{it}, RD_{it}, Publicity_{it}, ER_{it}, \pi_{it}) \tag{3.33}$$

Assuming $g(\bullet)$ *linear*, then,

$$\hat{u}_{it} = \delta_0 + \delta_1 Export_{it} + \delta_2 RD_{it} + \delta_3 Publicity_{it} + \delta_4 ER_{it} + \delta_5 \pi_{it} + W_{it} \tag{3.34}$$

Where *dummy variables* are: $Export_{it}$, 1 if firm i exports at time t, 0 otherwise; RD_{it}, 1 if firm i invests in Research and Development at time t, 0 otherwise and $Publicity_{it}$, 1 if firm i invests in publicity at time t, 0 otherwise. On the other hand, *decision variables* are: ER_{it}, which represents exchange rate (Mexican pesos per dollar), and π_{it}, which represents the annual inflation rate (based in Producer Price Index). W_{it} is the OLS error term, which must satisfied the classical assumptions.

The first three dummy variables include those qualitative aspects that could have influence in the efficiency of the firms. For example, one could expect that those export oriented firms being more efficient than those whose behavior is inward looking.

Similar analysis can be made in the R&D case. At the same time, publicity was included in order to capture behavior of those firms related with the competition, and how this kind of investments affect, or not, the efficiency; we expect that the more aggressive in publicity investment, the more efficiency the firm would have; this is not necessarily true, but could be understood as a signal of efficiency: a poor firm will not invest in publicity. Finally, exchange rate and inflation rate were included in order to capture their influence in the efficiency as indicator of "external influences" in the internal decisions of the firms.

3.3 The Mexican Manufacturing Sector Data Set

The data set used here has been obtained from the Annual Industrial Survey (AIS) applied to the Mexican manufacturing sector (sample). The full data set observes 6867 firms; nonetheless, we discarded those data which were uncompleted and unable to be analyzed because of the lack of information. Thus, we have in hand a panel-data sample with 4348 manufacturing firms followed 8 periods (1994-2001), and distributed in 9 subsectors (See appendix). Descriptive statistics for the data used in this study are given in Table 1:

Table 1. Descriptive Statistics

Variables	Means	Standard Deviation	Description
Production (Y)	218699.00	1132205.00	Thousands of Mexican Pesos
Labor (L)	21978.23	60315.36	Thousands of Mexican Pesos
Capital (K)	46482.96	219831.00	Thousands of Mexican Pesos
Other Inputs (O)	152382.20	816043.80	Thousands of Mexican Pesos
Export			1 if exports 0 otherwise
RD			1 if firm spend in R&D 0 otherwise
Publicity			1 if firm spend in Publicity 0 otherwise
Exchange Rate (ER)	7.968588	2.01293	Mexican Pesos / Dollars
Inflation ()	17.72844	12.62997	% (From Price Productor Index)

* Lowercase: y, l, k and o, are logs of its corresponding uppercase.

Behavior of the manufacturing sector, export oriented, shows three different stages as could be seen in Figure 1: first, the number of exporting firms raised 56 percent since NAFTA, 1994 to 1997 (crisis period). The second period shows an important failure (-10 percent) only in one year (1997-1998). Since then, the number of export manufacturing firms has been relatively constant (1998-2001). Additionally, Figure 2, shows that the exporting sector share of total production value at 1994 was 57 percent and grew gradually until 1997 when achieved its maximum value: 77 percent. Since then, there have not had significant changes and its share on total production value is around 74 percent.

Figure 1: Manufacturing Export Sector Evolution (1994-2001)

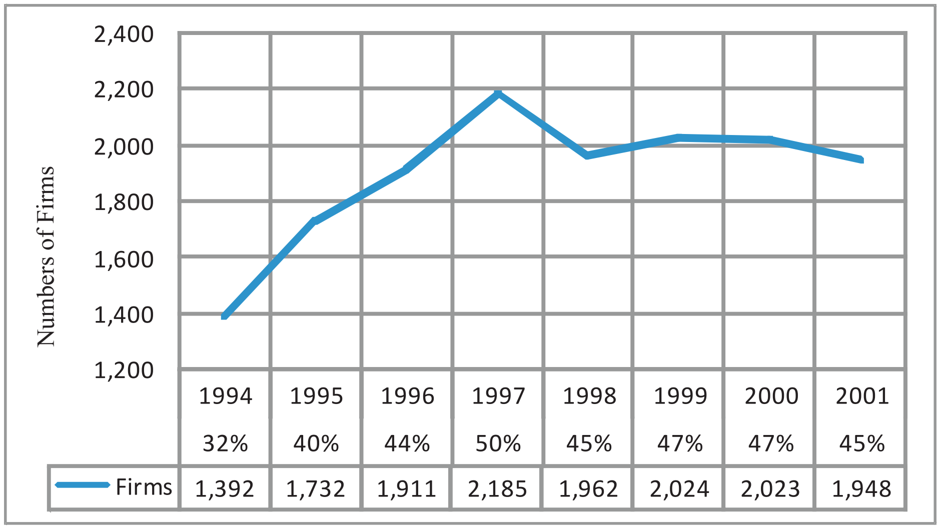
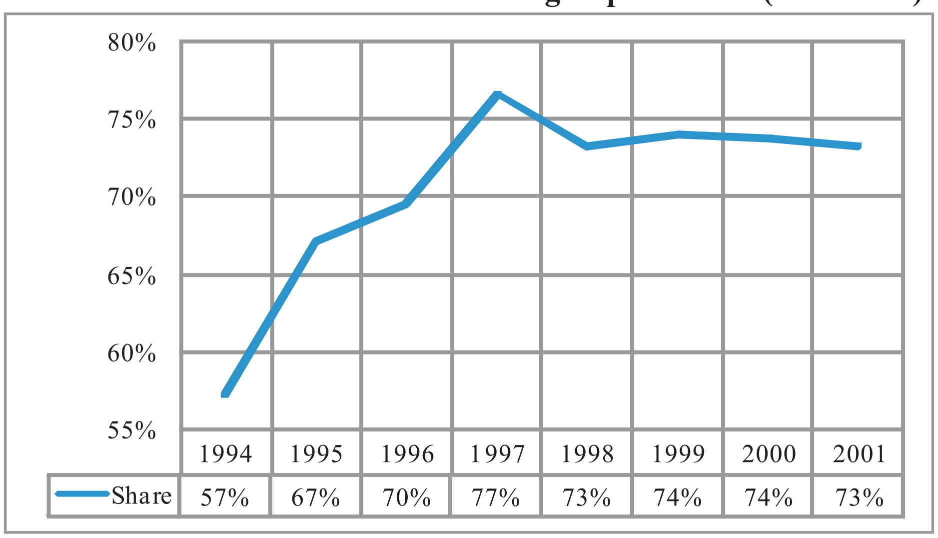


Figure 2:
Total Production Value Manufacturing Export Share (1994-2001)



In order to capture these changes, the empirical strategy will be the next: 1) we will study the whole period: 1994-2001, additionally, as the manufacturing sector seems to have had a structural change, we will try to capture it dividing main period in two parts: 2) 1994-1997 (Mexican Crisis period); 3) 1998-2001 (Transitional period).

3.4 Empirical Results

The stochastic frontier panel data model was estimated, as was mentioned, using the maximum likelihood method which standard errors were corrected using bootstrap techniques (1000 replications). The estimated coefficients of Equation (3.31) in its two models, time-invariant and time varying, are presented in Table 2.

3.4.1 Period 1994-2001: Production Function Results

Table 2. Production Function Estimated Coefficients (1994-2001)

	Time Invariant Model		Time Variant Model	
	Coeff.	t-value	Coeff.	t-value
Cosntant	2.4273	0.43	1.9855	32.48
l=ln(L)	0.1136	41.15	0.1255	44.31
k=ln(K)	0.0194	10.63	0.0224	12.31
o=ln(O)	0.8603	360.46	0.8623	362.63

This is the calculation of the Cobb-Douglas coefficient. Then, the constant term in a Cobb-Douglas function represents the *total factor productivity*, and is a variable which accounts for effects in total output not caused by inputs. In time invariant model, the constant term was no significant, this means that the the effects in total output depends only of the inputs. On the contrary, in time variant model the constant term was significant, which means that there is effects in total output driven by other reasons and not caused by inputs, which make sense since we allowed technical efficiency changes in each period (time-varying model).

All coefficient inputs (time invariant and time variant models) are positive and statistically significant at 1-percent level; results are very similar between two models.

Eventhough constant term in the time invariant model was no significant, we can estimate u_i using the steps described in section 2. Then, having in hand estimated inefficiencies for both models, u_i and u_{it}, we can compare them. The correspondence between both sets of estimation draws attention: Table 3 shows that pairwise correlation is near to 1 (0.9148).

Table 3. Analysis of Estimated Technical Inefficiencies

	Time Invariant	Time Variant
Mean	1.7310	1.5080
Standar Dev.	0.2063	0.1907
Correlation	0.9148	

We can rank, in terms of inefficiency[11], producers of the Mexican manufacturing sector using both models: time-invariant and time-variant. We observe that time-invariant model reproduces, in more than one case, the ranking of time-variant model as can be seen in Figure 3. However, distribution of inefficiency is different between time-invariant and time-variant models as can be seen in inefficiency kernel[12] estimated distribution (Figure 4). For instance, mean and standard deviation is greater in time-invariant case, $\hat{u}_i$, compared with time-variant $\hat{u}_{it}$ case[13]. This implies less variability in time-variant model of inefficiency. Eventhough this difference between models exists, the ranking made for both models is almost the same as it will be seen

[11]Which is analogous if we express it in terms of efficiency, because we use a monotone transformation of error terms $(\hat{TE_i}) = exp(-\hat{u}_i)$. The scatter plot of $(\hat{TE_i}) = exp(-\hat{u}_i)$ versus $(\hat{TE_i}) = exp(-\hat{u}_{it})$ is very similar to Figure 3.

[12]We use Epanechnikov kernel function:

$$\hat{f} = \hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\Big(\frac{x_i - x}{h}\Big)$$

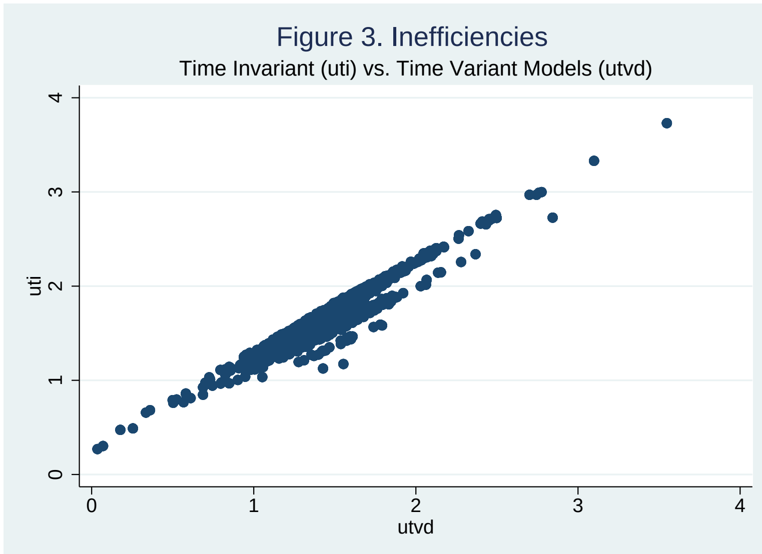
where

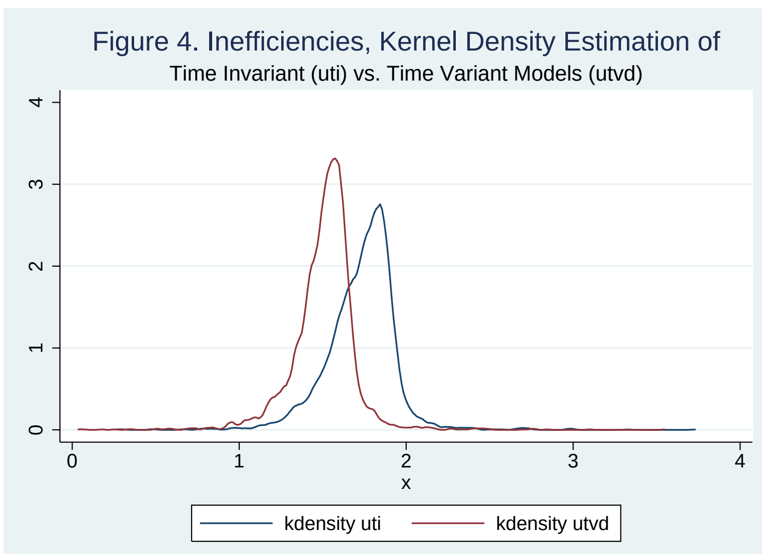
$$K(u) = \frac{3}{4}(1-u^2)\mathbf{1}_{\{|u|\leq 1\}},$$

where $\mathbf{1}_{\{A\}} = 1$ if A holds; 0, otherwise. Optimal bandwidth was used. See Pagan and Ullah (1999), pg 28.

[13]*Time-invariant* mean and standard deviation were 1.73 and 0.21 respectively; *time-variant* mean and standard deviation were in turn 1.51 and 0.19.

below.





3.4.2 Period 1994-2001: Sources of Inefficiency

Table 4 presents a second step analysis in order to identify sources of inefficiencies from the two models. Since u_i is given in proportional terms, the absolute magnitudes of the coefficients give the proportional impacts (Greene, 2002). Results in both models, suggest that exports and publicity are significant in explaining variation in efficiency. Exports impulse the efficiency (negative sign)[14], and publicity reduces it (positive sign), maybe because of its impact in the costs. In particular, R&D has ambiguous results: in the expected direction, time invariant model shows that R&D is a source of efficiency (negative sign and significant). On the contrary, in time variant model, R&D is not significant. Figure 1A in appendix could help us to explain the problem: it shows us the number of firms that invested on R&D. It should be noticed that from 1994 to 1997, the percentage of total number of firms that invested in R&D grew from 9 to 23%; since then, this percentage felt until it reached 16 percent and stayed. It seems that time-variant model capture this behavior which explains that R&D was not statistically significant. Finally, neither exchange rate nor inflation[15] has statistic significance. It means that inefficiency depends on firm's internal structure and decisions rather than external influences such as prices or exchange rate.

Table 4. Second Step Regression Results

	Time Invariant		Time Variant	
	Coeff.	t-value*	Coeff.	t-value*
Constant	1.7343	774.78	1.5046	732.34
Export	-0.0195	-9.03	-0.0088	-4.2
R&D	-0.0147	-4.76	-0.0032	-1.14
Publicity	0.0115	4.87	0.0118	5.47
ER	0.0005	0.95	0.0001	0.27
πi	0.00001	0.938	0.00001	0.07

*Standard Errors Bootstrap Corrected.

3.4.3 Period 1994-2001: Firms Performance

Using $\hat{TE_i}$ in both models, time-invariant and time-variant, we can compare performance between firms, in terms of the "benchmark" explained in section 2. Table 5, shows the ten most efficient firms in the manufacturing sector and their levels of

[14]Increases in u_i imply lower efficiency (Greene, 2002).
[15]Producer Price Index percentage change is used in this case as a measure of inflation rate.

technical efficiency over the period 1994-2001 using both models: time-invariant and time-variant.

Table 5. Firms Rank Based on Stochastic Frontier Model with Time-Invariant and Time-Variant assumptions

#Id	Time Invariant	Time Varying	SIC Classif.	Description
1654	1	1	351214	Industrial Gas Manufacturer
376	2	2	314002	Cigarette Manufacture
1510	3	3	351214	Industrial Gas Manufacturer
775	4	4	351214	Industrial Gas Manufacturer
1434	5	5	351214	Industrial Gas Manufacturer
1362	6	6	351214	Industrial Gas Manufacturer
477	7	8	314002	Cigarette Manufacture
1067	8	10	351214	Industrial Gas Manufacturer
3704	9	7	355003	Natural or Synthetic Pieces or Articles Rubber Manufacture
1536	10	9	382106	Joint and Repairer of Machinery and Equipment for other Specific Industries

In general, the two models fit very well and almost coincide; in fact, the first six firms have the same rank in both models; on what remains, they have minimum differences. Interestingly, six firms of the top ten, including the first one, were classified as 351214 (Industrial gas manufacture) and the total number of firms in this classification represents, surprisingly, only 0.34% of the whole manufacturing sector.

Those that are ranked as 2nd and 7th (2nd and 8th, respectively, considering time-variant model), were classified as 314002 (Cigarettes manufacture). The 9th (or 7th in time-variant case) was classified as 355003 (Natural or synthetic pieces or rubber manufacture articles) and, finally, 10th (or 9th in time-variant model) was classified as 382106 (Manufacture, joint and repair of machinery and equipment for other specific industries).

It seems that time-invariant model underestimate potential efficiency of the firms because of the contrast that can be seen between models. For example, the most efficient firm (1654) worked at 76% of its potential output in the time-invariant model, whereas worked near to 97% of its potential output in time-variant model (see Table 5.1). Although this happened, the important thing is that the ranking fitted well between models as was said before.

Table 5.1. Firms Performance Indices Based on Stochastic Frontier Model with Time-Invariant and Time-Variant Assumptions

#Id	Time Invariant	Time Varying
1654	0.7646	0.9664
376	0.7402	0.9355
1510	0.6237	0.8450
775	0.6141	0.7845
1434	0.5190	0.7261
1362	0.5061	0.7088
477	0.4680	0.6178
1067	0.4650	0.5812
3704	0.4547	0.6199
1536	0.4518	0.6055

Finally, following the time-invariant model, in average, manufacturing sector, as a whole, is working at 18 percent of its potential product (1994-2001); on the other hand, considering the time-variant version of our model, which includes certain dynamic behavior, we observed a falling in efficiency. Indeed, should be highlight that in 1994 the manufacturing sector as a whole was working at almost 24 percent of its potential product; on the contrary, in 2001 was working at 22 percent (see Figure 4A in Appendix).

3.4.4 Period 1994-1997 vs. 1998-2001: Production Function

Now, in order to capture structural change, if there exists, analysis will be done by dividing the whole period in two parts: 1994-1997 (Crisis) and 1998-2001 (Transition). Considering the length of both symmetric periods, 4 years each, we are able to use time-invariant assumption, explained in Section 2 above, for modeling technical efficiency.

Table 6. Production Function Coefficients

	1994-1997		1998-2001	
	Coeff.	t-value	Coeff.	t-value
Cosntant	2.4794	0.25	2.4886	0.33
l=ln(L)	0.1464	39.83	0.0978	25.76
k=ln(K)	0.0287	11.70	0.0222	8.49
o=ln(O)	0.8329	274.70	0.8832	265.08

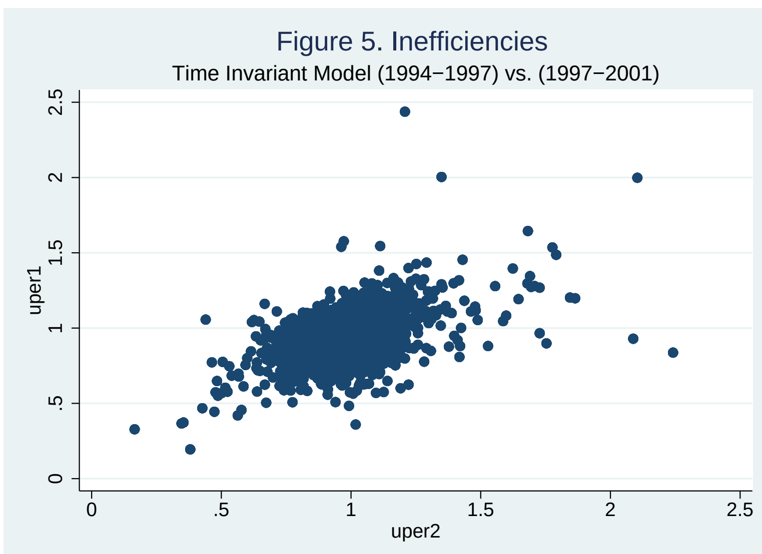
Given the time invariant assumption, in both cases the constant was no significant. In both periods, capital (k) and other inputs (o) are relatively similar: 0.0287 vs. 0.222 and 0.8329 vs. 0.8832 respectively; but, that is not the case of the labor force which was 0.1464 in the first four years and 0.0978 in the next period.

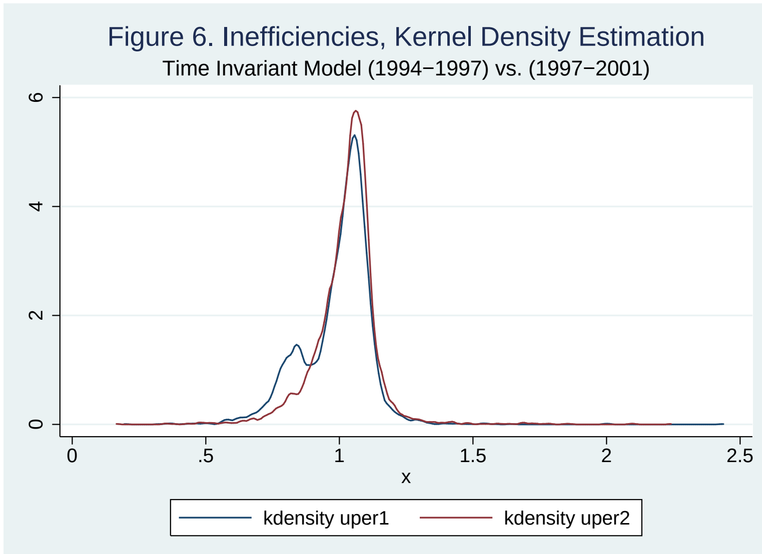
Taking account the fact that our model was expressed in logs, then, coefficients of the production function can be understood as elasticities. The relevant one is β_1: the elasticity of the production with respect to the labor force. Indeed, this coefficient felt from 1994–1997 to 1998–2001, i.e., in the first period, *ceteris paribus*, a 1% increment in labor would lead approximately 0.14% increase in output; in the second period, the same increase in labor (1%), would lead only 0.09% increase in output, less than in the first period. It means that productivity of the labor decayed between periods.

Inefficiency measure, u_i, is notably different between two periods as can be seen in Figure 5. There is not a clear pattern in the scatter plot which means that inefficiency changed in time, eventhough the distribution of inefficiency in both periods was relatively similar (see Figure 6). Another component that gives us information about the existence of structural change is the relatively low pairwise correlation coefficient between inefficiency estimated in both periods: 0.4685.

Table 7. Analysis of Estimated Technical Inefficiencies

	1994-1997	1998-2001
Mean	0.9927	1.0264
Standar Dev.	0.1291	0.1197
Correlation	0.4685	





3.4.5 Period 1994-1997 vs. 1998-2001: Sources of Inefficiency

Results are shown in Table 8.

Table 8. Second Step Regression Results

	1994-1997		1998-2001	
	Coeff.	t-value*	Coeff.	t-value*
Constant	0.9884	552.79	1.0267	520.87
Export	-0.0075	-3.84	-0.0080	-4.34
R&D	-0.0135	-4.5	0.0024	0.95
Publicity	0.0147	7.46	0.0045	2.18
ER	0.0006	0.6	-0.0003	-0.03
πi	-0.00004	-0.33	0.000005	0.01

*Standard Errors Bootstrap Corrected.

In both periods the constant term is positive and significant. In both cases, export is a source of efficiency (negative sign). On the other hand, as was seen above, spending on publicity reduces the efficiency of the firm in both periods. As in the whole period, neither exchange rate (ER) nor inflation rate (π), as decision variables, were significant. Figure 2A and 3A in the appendix, could help to understand why: in the first one, exchange rate (ER) shows only one important change (1994/12-1995/01) which is the crisis period; since then, eventhough ER rises, monetary policy in Mexico seems to have been efficient and ER was stabilized. In the second one, the inflation rate was drawn. In the first period (1994-1997), inflation rate dramatically arouse, even more than 50 percent; since then, fall gradually until certain stabilization. Then, instability of those variables in the first period could have been the reason because of what firms did not take account them as a decision variables.

Finally, the main source that could explain the structural change is R&D variable. Indeed, in the period 1994-1997 R&D, as expected, was negative and significant which means that a major efficiency was observed. But in period 1998-2001 was not significant which means that firms lost the confidence in R&D as a source of efficiency (see Figure 1A in the appendix).

3.4.6 Period 1994-1997 vs. 1998-2001: Firms Performance

Table 9, shows firms ranking comparison between periods.

Table 9 Firms Rank and Performance Indices Based on the Stochastic Frontier Model (1994-1997 vs. 1998-2001)

iid	Rank 1994-1997	Efficiency	Rank 1998-2001	Efficiency	subsector
376	1	0.8247	4	0.6846	31
1654	2	0.7217	1	0.8484	35
2298	3	0.6994	1705	0.3618	31
1510	4	0.6943	2	0.7080	35
775	5	0.6900	3	0.7026	35
1067	6	0.6582	18	0.5701	35
1362	7	0.6423	8	0.6238	35
477	8	0.6347	21	0.5619	31
1434	9	0.6275	5	0.6535	35
162	10	0.6177	1280	0.3712	32
3704	20	0.5644	9	0.6210	35
2361	256	0.4626	7	0.6298	35
2310	2310	0.5237	10	0.6174	35
3645	2885	0.3483	6	0.6451	33

In general, the top ten firms of the first period (1994-1997) worked at 68 percent of their capability (in average); in the second period, 1998-2001, there is a slight losing of technical efficiency because, in average, the firms worked at 67 percent of their capability.

Behavior of firms in terms of efficiency measure, reveals that some enterprises have been consistent, but, at the same time, it is possible to detect some winners and some losers. Then, we can define as "consistent" firms those that belong to top ten in the first period and in the second period too (eventhough, they were not in the same ranking); we can define as "winners" those firms that do not belong to top ten in first period but in the second one they are; and, finally, we defined: "losers" to those firms that were in top ten in the first period and were not in the second one.

Taking account this classification, we have 6 "consistent" firms which, surprisingly, are the same first six firms ranked in the Table 5 of both models that were seen before, including 5 firms (the first one too) that are classified as 351214 (Industrial gas manufacture) and, ranking in 2nd place, classified as 314002 (Cigarettes manufacture). The most efficient firm worked at 82 percent of its potential output[16]. "Winners" are 4, but the relevant one is identified as 3645 (see first column of Table 9) which in the first period was placed in 2885th; in the second period, this firm was ranked in the 6th place and is classified as 332001 (Manufacture and repair of furniture, wood mainly). This firm worked at 35 percent in the first period and at 65 percent at the second period. Finally, there are 4 "losers", but the relevant one is the firm number 2298 (see Table 9) which is classified as 312200 (Food preparation and mixture for animals) that felt from 3rd place to 1705th place; moreover, this firm worked at 70% of its potential output in the first period and at 36%, in the second one.

The manufacturing sector, as a whole, in the first period (1994-1997), according with this model (time-invariant), was working at 37 percent of its capacity; in the second period (1998-2001) felt and worked at 36% of its potential output.

Then, the 1995's Mexican crisis was not itself the main cause of lacking development of the manufacturing sector, but the absence of R&D investments; weak capability of adaptation for "fighting" successfully against the foreign firms; and lacking development of efficiency in the new scenario. Despite this facts, there were some firms that "survived" to the openness and were "consistent"; whereas other firms were "winners" and other firms were "losers" in the process.

3.5 R&D Strategic Interaction

In the context of NAFTA, we would expect that interaction between firms have had place. Considering, for example, the survivor analysis, we expect that firms increase its investment in R&D, taking account that they believe that other firms like them would do the same thing. In order to capture interaction behavioral between firms,

[16]Should be remembered that time-invariant model seems to underestimate the potential efficiency of the firms. And we should remember that this efficiency is related to the "benchmark" firm(s).

in particular strategic interaction between the two most important firms of each six digit manufacturing sector in R&D decisions[17], we assume that this two firms interact under a Cournot competition; in other words, each firm should decide the quantity given the expected action to the other firm and the demand structure.

In the simplest model of Cournot competition, we assume that the demand structure is given and unknown by the econometrician. Then, each firm maximize its profits, and price is a commonly known decreasing function of total output. We assume that each firm has a cost function $c(q_i)$, which is a "marginal constant function". This function is an increasing function of the q_i. Market price is set at a level such that demand being equal to the quantity produced by two firms (duopoly). Assuming Nash equilibrium, we can conclude that $q_1 = q_1(q_2)$. Symmetric result is obtained for q_2. Given the equilibrium quantity, firms observe their equilibrium profits, Π_i^*.

3.5.1 Assumptions

A1. Each of two firm in six digit level produce an homogeneous product. Firms do not cooperate, and because of its size have market power, and compete in quantities, choosing quantities simultaneously. The econometrician does not observe this quantities, neither firm's equilibrium profits Π_i^* (latent variable). Player 1 and player 2 are distinguished each other for the size (which is determined by its sales share): player 1 is the largest firm in each couple of players.

A2. It is assumed that Π_i^*, derived from Nash equilibrium, can be expressed as follows:

$$\Pi_i^* = \boldsymbol{\beta}'\mathbf{X}_i + \varepsilon_i \tag{3.35}$$

for *i=1,2*.

Following Aradillas-Lopez, (2003),

[17] By "most important" we understand those firms whose market share set them in the first and second place of all this six digit industry.

A3. $\mathbf{X}_1 \in \mathbb{R}^k$ and $\mathbf{X}_2 \in \mathbb{R}^k$ are independent draws from the same distribution with (joint) cdf given by *$F(\boldsymbol{x})$*, and corresponding pdf given by *$dF(\boldsymbol{x})$*

A4. $\varepsilon_1 \in \mathbb{R}$ and $\varepsilon_2 \in \mathbb{R}$ are independent draws from the same distribution with cdf given by $G(\epsilon)$.

A5. ε_i is independent from $\mathbf{X}_i$ for $i \in \{1, 2\}$

A6. *At the time the game is played*, the realizations of $(\mathbf{X}_1, \varepsilon_1)$ and $(\mathbf{X}_2, \varepsilon_2)$ are privately known by players 1 and 2 respectively.

A7. Distributions $(F(\boldsymbol{x}), G(\epsilon))$ are known by both players.

Now, suppose that some time after the game was played, the econometrician have access to M outcomes of the players and the following is true:

B1.-Assumptions (A1-A7) were satisfied when the game was played by each of the N pairs of players.

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

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

B4.-The distribution $G(\epsilon)$ 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 of $F(\boldsymbol{x})$. We only assume that this distribution does not depend on any of the payoff parameters, beliefs or the unknown parameters of $G(\epsilon)$.

3.5.2 Decision rule

Now, let us define decision rule. There are two kind of actions that players can choose in this model: "to be aggressive" or "not" in the investment of R&D sense. A firm will be "aggressive" ($y_i = 1$) if $\mathbf{1}[\Pi_i^* > 0]$, where $\mathbf{1}[A]$ is the indicator function: equal to 1 if the event A is true, zero, otherwise.

3.5.3 The Model

Under this criteria, we are trying to determine the probability of being aggressive ($y_i = 1$), given the characteristics of the firms, i.e., $Pr(y_i = 1|\mathbf{x})$.

ε-distribution

As was said, we assumed that ε_i is orthogonal to $\mathbf{X}_i$. In order to be parsimonious and without losing of generality, we suppose that ε_i adopt a logistic distribution.

$$\Lambda(\epsilon) \equiv \frac{e^{\epsilon}}{1 + e^{\epsilon}} \tag{3.36}$$

Then, following Wooldridge (2001) pg. 457-469,

$$\begin{aligned} Pr(y_i = 1|\mathbf{x}) &= Pr(\Pi_i^* = \boldsymbol{\beta}'\mathbf{X}_i + \varepsilon_i > 0|\mathbf{x}) \\ &= Pr(\varepsilon_i > -\boldsymbol{\beta}'\mathbf{X}_i|\mathbf{x}) \\ &= \Lambda(\boldsymbol{\beta}'\mathbf{X}_i) \end{aligned} \tag{3.37}$$

Solving (3.37) by maximum likelihood methods, we can find the betas.

3.5.4 Empirical application

Let be $y_i = 1$ if there is a positive increment of R&D investment between period t and t+1, $y_i = 0$ otherwise. Periods taken account in this model were $t \in \{94, 96, 98\}$. The difference between years was made in order to mitigate the time influence. M=235, which means that there were 470 firms. That is why were used Standard Errors corrected by bootstrap (1000 replications).

Variables

Dependent variable: this is a dichotomous variable y_i, which values are 1, if $i's$ firm is aggressive, i.e., if $i's$ firm increase its R&D investment between t and t+1; zero otherwise. The firm taken account in this case was the smallest one of the two firms considered here, in each six digit industry.

Other firm's actions: there is a dichotomous variable too, y_{-i}, which criteria is the same that was taken in the dependent variable. These are the biggest firms, according with the size criteria, which was constructed using the market share in each sector.

Herfindhal index: market structure could have influence in the competence between the two most important firms in each six digit manufacturing sector. Then, the concentration of the industry influence in the decisions of the firms is capture by Herfindhal index[18], H_i.

Price producer index: evidently, changes in the prices affronted by the producer could affect his R&D investment decisions. We expect a negative influence between price producer index, P_i, and R&D investment decisions.

$$\mathbf{X}_i = \{y_{-i}, H_i, P_i\}_{i=1}^{M} \tag{3.38}$$

where M=235.

3.5.5 Results

Solving (3.37), and using (3.38), we have:

[18] In general, manufacturing sector is not concentrated. On average, Herfindhal index reach the value of 0.13.

Table 1. Estimation Results

(Standard Errors[19] in parentheses)

y_{-i}	1.8049*
	(0.4609)
H_i	2.6404*
	(1.3866)
P_i	-0.0153*
	(0.0021)

(∗) Statistically significant at a 5% level.

This result shows, that the decision made by the largest firm, i.e., if there it is aggressive in the R&D investment sense, impulses the smaller one to be aggressive, and this can be seen in the sign of the coefficient which is positive and statistically significant. On the other hand, the more concentrated the industry is, the more aggressive the smaller one tends to be; this can be explained evoking the survivor analysis. In a competitive context (perfect competition), firms can survive with higher probability, without necessity of being aggressive in the R&D investment; on the contrary, in a concentrated industry firms need to invest in R&D in order to compete and survive, that is why the coefficient sign is positive and statistically significant. Finally, as was expected, an increase of the general level of prices that producers affront, tends to restrain the R&D investment impulse of the smaller firm: this can be seen in the negative sign of the coefficient, which is statistically significant too.

Considering the structure of the model, "logit model", we can calculate marginal effects ($\partial y_i/\partial \mathbf{X}_i$) which are presented in table 2.

[19]Bootstrap corrected using 1000 replications

Table 2. Marginal Effects

(Standard Errors in parentheses)

y_{-i}	0.2472*
	(0.1011)
H_i	0.2208*
	(0.1127)
P_i	-.0013*
	(0.0002)

(∗) Statistically significant at a 5% level.

Sign of marginal effects have been inherited of the coefficients signs. The probability of smaller firm tending to be aggressive is 0.25. In other words, for each large firm that increase its R&D investment, 1/4 of firms will increase its R&D investment too. Similar analysis can be made with the other coefficients.

3.6 Conclusions

Manufacturing sector in Mexico is not homogenous. This assertion is confirmed by the models presented in this paper: each firm observes different level of efficiency. For instance, the worst firm in *time-invariant* model (1994-2001) was working at 2.4 percent of its capacity (2.8 percent in *time-variant* model), compared with the "benchmark" firm[20]; the best one (which is the same in two models), classified as 351214 (Industrial gas manufacture), was working at 76 percent of its capacity (97 percent in time-variant model). In average, manufacturing sector was working at 18 percent of its potential product (23 percent in the time-variant case), understanding

[20]The worst firm is the same in these two models and is classified as 311301 (Preparation and packaging of fruits and vegetables).

"potential product" in comparison with the best firm(s) performance to the manufacturing sector. At the same time, should be highlight that in 1994 the manufacturing sector as a whole was working at almost 24 percent of its potential product but in 2001 at 22 percent, which means a loss of its capacity.

On the other hand, the second part of this paper shows the existence of structural change. Indeed, the model in which a partition of the whole period was made: 1994-1997 and 1998-2001, shows how some firms were "consistent" keeping its performance in top ten ranking; how an other firms were "winners" remaining in top ten ranking in the second lapse; and finally, how some firms were "losers", i.e., those firms that in the first period were in top ten and not in the second one. Then, eventhough the crisis period (1995), the second lapse (1998-2001) shows certain stability, however, there was a lost of the potential capabilities of the manufacturing sector, maybe because of the openness and the entrance of foreign manufacturing products since NAFTA.

Finally, calls the attention that "industrial gas manufacture" had 6 firms in the top ten ranking (in both models) which means that NAFTA seems have not had an important effect in another manufacturing subsectors; on the contrary, seems that had a harmful effect in whole manufacturing sector given the loss of competence observed.

It seems that manufacturing sector was resistent to the openness in those cases in which natural resources give some advantages (gas resources, tobacco (natural conditions), etc.); but, its not the case of those firms in which was necessary to compete ("fight", in the IO jargon). This firms should have been more efficient.

In terms of R&D decisions we can say that, under a Cournot competition equilibrium, firms care about other actions, in particular, smaller firms tend to be aggressive if the biggest one of the six digit manufacture sector, is aggressive. Actually, the probability that small firms would tend to be aggressive is 0.25. NAFTA created competence between the two most important firms in each six digit manufacturing sector.

Appendix C3

Table 1A. Manufacturing Sector Classification
31 - Manufacture of Food, Beverages and Tobacco
32 - Textile, Wearing Apparel and Leather Industries
33 - Manufacture of Wood and Wood Products, Including Furniture
34 - Manufacture of Paper and Paper Products, Printing and Publishing
35 - Manufacture of Chemicals and Chemical, Petroleum, Coal, Rubber and Plastic Products
36 - Manufacture of Non-Metallic Mineral Products, except Products of Petroleum and Coal
37 - Basic Metal Industries
38 - Manufacture of Fabricated Metal Products, Machinery and Equipment
39 - Other Manufacturing Industries

It was used the Mexican Classification of Activities and Products (MCAP) in its 1994 version, which is compatible with Uniform International Industrial Classification (UIIC) at four digit level.

Table 2A.Firms in each manufacturing subsector		
Subsector	**Firms**	**%**
31	819	18.84
32	664	15.27
33	160	3.68
34	357	8.21
35	901	20.72
36	285	6.55
37	101	2.32
38	1,010	23.23
39	51	1.17
Total	4,348	100

Figure 1A. Percentage of Total of Firms that Invested in R&D (1994-2001)

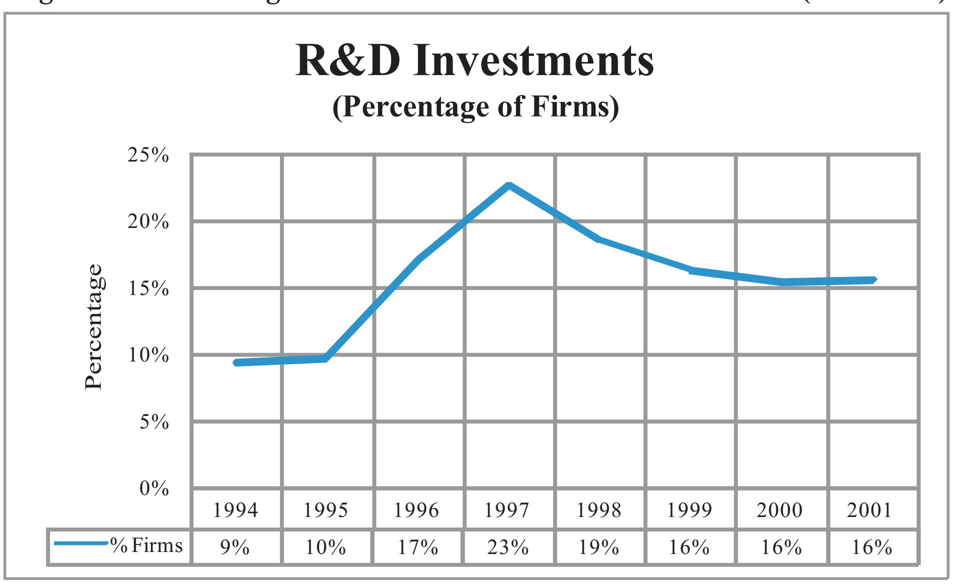


Figure 2A

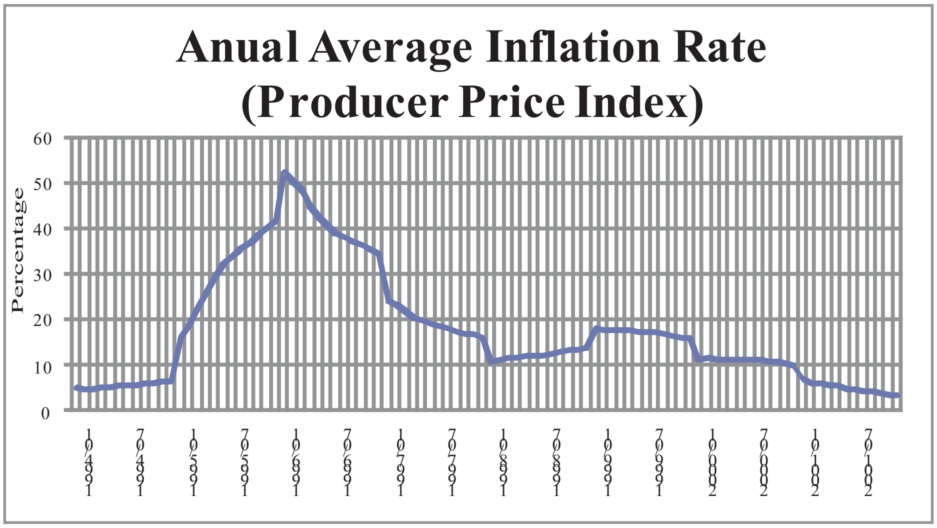


Figure 3A

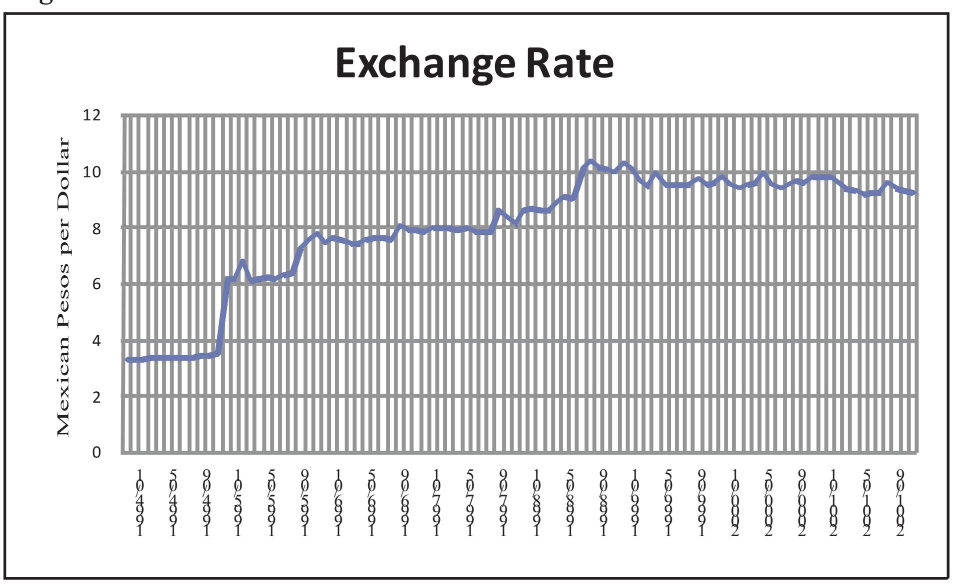
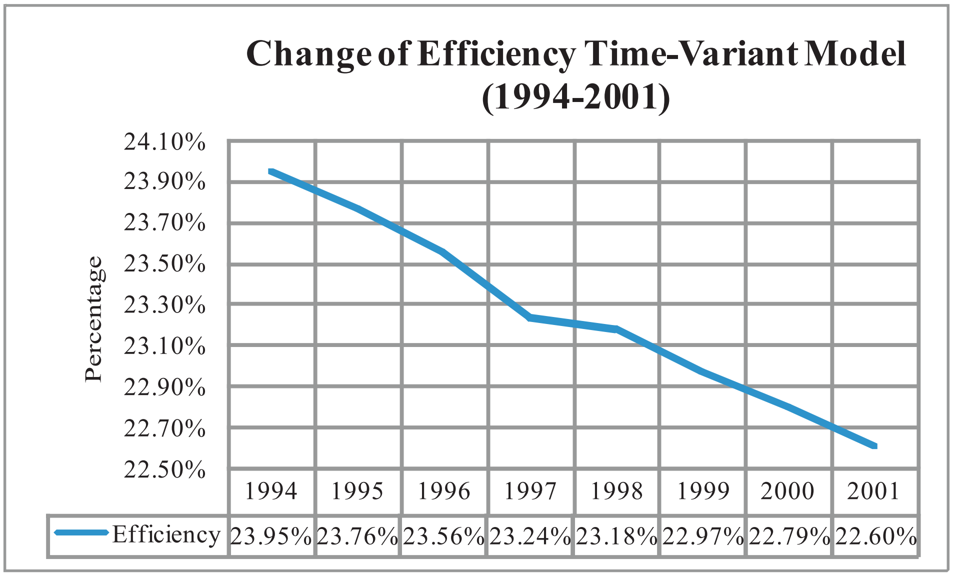


Figure 4A


Change of Efficiency Time-Variant Model
(1994-2001)
Percentage
24.10%
23.90%
23.70%
23.50%
23.30%
23.10%
22.90%
22.70%
22.50%
1994
1995
1996
1997
1998
1999
2000
2001
Efficiency
23.95%
23.76%
23.56%
23.24%
23.18%
22.97%
22.79%
22.60%

Bibliography

[1] Aguirregabiria, V., and Mira, P. (2002): *Sequential Simulated-Based Estimation of Dynamic Discrete Games*, Boston University Working Paper.

[2] Aguirregabiria, V. (2005): *Another Look at the Identification of Dynamic Discrete Decision Processes* Boston University, Department of Economics.

[3] Aigner, D., K. Lovell, K. and P. Schmidt. *Formulation and Estimation of Stochastic Frontier Function Models*, Journal of Econometrics, 6, 1977,pg. 21-37.

[4] Amemiya, Takeshi, *Advanced Econometrics*, Harvard University Press, Cambridge, Massachusetts, 1985.

[5] Aradillas-Lopez, A. (2003). *Empirical Likelihood Estimation of Interactions-Based Models with Asymmetric Information.* Unpublished Manuscript, Princeton University.

[6] Aradillas-Lopez, A. (2008). *Semiparametric Estimation of a Simultaneous Game with Incomplete Information.* Unpublished Manuscript, Princeton University.

[7] Badunenko, Oleg, Fritsch, Michael, and Stephan, Andreas, 2006. *What Determines the Technical Efficiency of a Firm? The Importance of Industry, Location, and Size*; Jenaer Schriften zur Wirtschaftswissenschaft from Friedrich-Schiller-Universitt Jena, Wirtschaftswissenschaftliche Fakultt.

[8] Bajari, P., Benkard, C.L., and Levin, J. (2003): *Estimating Dynamic Models of Imperfect Competition*, Stanford and Duke University Working Paper.

[9] Bajari, P. and Hong, H. (2005), *Semiparametric Estimation of a Dynamic Game of Incomplete Information*. Technical Working Paper, 320. NBER 2006.

[10] BANSEFI (2004). *Encuesta Basal sobre el Ahorro, Crédito y Microfinanzas Rurales*. Reporte Metodológico.

[11] Baltagi, Radi H. (2005) *Econometric Analysis of Panel Data* Ed. Jhon Wiley and Sons.

[12] Battese, G. and T. Coelli (1992), *Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India*, Journal of Productivity Analysis, 3, 1, pp. 153-169.

[13] Battese, G. E., Coelli, T., (1995). *A model for technical inefficiency effects in a stochastic frontier production function for panel data*. Empirical Economics 20, 325-332.

[14] Bjorn, P.A., and Q. Vuong (1984). *Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation* SSWP No. 537, Caltech.

[15] Bjorn, P.A., and Q. Vuong (1985). *Econometric Model of a Stackelberg Game with an Application to Labor Forece Participation* SSWP No. 577, Caltech.

[16] Bresnahan, T., and P. Reiss (1991): *Empirical Models of Discrete Games. Journal of Econometrics*, 48, 57 - 81.

[17] 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.

[18] Blundell, R. and Smith J. R. (1993). *Simultaneous Microeconometric Models with Censored or Qualitative Dependent Variables*. In G. S. Maddala, C. R. Rao and H. D. Vinod (eds.) *Handbook of Statistics*, Vol 11 (New York: North Holland) 117-143.

[19] Blundell, R. and Smith J. R. (1994). *Coherency and Estimation in Simultaneous Models with Censored or Qualitative Dependent Variables. Journal of Econometrics*, **48**, 57-81.

[20] Calderon M. A. and Voicu A., 2004. *Total Factor Productivity Growth and Job Turnover in Mexican Manufacturing Plants in the 1990s*. IZA DP. No. 993, Forschungsinstitut zur Zukunft der Arbeit Institute for the Study of Labor.

[21] Calderón-Madrid, A. (2007) *Output Fluctuations, Employment Protection Regulations and the Mobility of Workers Between Formal and Informal Job Status in Mexico*. Under Review Unpublished Manuscript.

[22] Caves, Richard E., 1989. *International differences in industrial organization*. In: Schmalensee, Richard, and Willig, Robert, (Eds.), Handbook of Industrial Organization, vol. 2. Elsevier, pp. 1225-1250.

[23] Coelli, Tim, D.S. Prasada Rao, and George E. Battese (2002) *An Introduction to Efficiency and Productivity Analysis.* London.Kluwer Academic Publishers, London.

[24] Cornwell, C., P. Schmidt and R. C. Sickles (1990), *Production frontiers with cross-sectional and time-series variation in efficiency levels*, Journal of Econometrics, 46, 185-200.

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

[26] Dixit, A.K. and R.S. Pindyck (1994). *'Investment Under Uncertainty'*. Princeton, NJ: Princeton University Press.

[27] Gong, Xiaodong and Arthur van Soest (2002). *Family Structure and Female Labor Supply in Mexico City. The Journal of Human Resources*, Vol. 37, No. 1, pp. 163-191. University of Wisconsin Press

[28] Greene, W., *A Gamma Distributed Stochastic Frontier Model*, Journal of Econometrics, 46, 1, (1990), pp. 141-164.

[29] Greene, William, (1997). F*rontier Production Functions*. In: Pesaran, M. Hashem, and Schmidt, Peter, (Eds.), Handbook of Applied Econometrics, vol. II. Blackwell Publishers, pp. 81-166.

[30] Greene, William, (2003). *Econometric Analysis*, Prentice Hall, New Jersey, 2003, 5th Ed.

[31] Greene, William, (2005). *Reconsidering Heterogeneity in Panel Data Estimators of the Stochastic Frontier Model*. Journal of Econometrics 126: 269-303.

[32] Greenstein, Theodore N. (1996). *Husbands' Participation in Domestic Labor: Interactive Effects of Wives' and Husbands' Gender Ideologies. Journal of Marriage and the Family*, Vol. 58, No. 3, pp. 585-595. National Council on Family Relations

[33] Hall and Vopel (1997). *'Innovation, market share and market value'* NBER.

[34] Hall, P. (1994) *Methodology and theory for the bootstrap*. Handbook of Econometrics, Volume IV, Edited by R.F. Engle and D.L. McFadden, Elsevier Science B.V.

[35] Hall (1999). *'Innovation and market value'* NBER Working Paper No. 6984

[36] Hayashi, Fumio, 2000. *Econometrics*, Princeton University Press; Princeton, New Jersey.

[37] 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.

[38] Heckman, J. (1978). *Dummy Endogenous Variables in a Simultaneous Equation System. Econometrica*, 46, 931-959.

[39] Heckman, J., and S. Navarro (2005): Dynamic Discrete Choice and Dynamic Treatment Effects, University.

[40] Hisao, Cheng (2003), *Analysis of Panel Data*, Econometric Society Monographs, Cambridge University Press. New York.

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

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

[43] INEGI, *Methodological Synthesis of Annual Industrial Survey*. 2005.

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

[45] Jenkins, M., P. Liu, D. McFadden, and R. Matzkin (2004): The Browser War: Econometric Analysis of Markov Perfect Equilibrium in Markets with Network Effects, UC Berkeley, working paper.

[46] Johnson R. A., James W. Evans and David W. Green (1999). *Some Bivariate Distributions for Modeling the Strength Properties of Lumber*. USDA.

[47] Keane, M. P. and K. I. Wolpin (1997, June). *The career decisions of young men.* Journal of Political Economy 105(3), 473.522.

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

[49] Kitamura, Y. (2006). *Empirical Likelihood Methods In Econometrics: Theory and Practice* Cowles Foundation Discussion Paper No. 1569.

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

[51] 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.

[52] Kooreman, P. (1994). *Estimation of Econometric Models of Some Discrete Games. Journal of Applied Econometrics*, Vol. 9, No. 3, (Jul. - Sep., 1994), pp. 255-268. John Wiley & Sons

[53] Kumbhakar, S. C. (1990), *Production frontiers, panel data, and time-varying technical inefficiency*, Journal of Econometrics, 46, 201-212.

[54] Kumbhakar, Subal and Knox, Lovell, C. A., (2000). *Stochastic Frontier Analysis*, Cambridge University Press, New York.

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

[56] Lee, Y. and P. Schmidt (1993), *A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency*, In H. Fried and K. Lovell, eds., The Measurement of Productive Efficiency: Techniques and Applications, Oxford University Press, New York, 1993.

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

[58] McFadden, D. (1974) *Conditional Logit Analysis of Qualitative Choice Behavior.* In *Frontiers of Econometrics.* Edited by P. Zarembka. New York: Academic Press.

[59] McFadden, D. (1976) "Quantal Choice Analysis: A Survey," Annals of Economic and Social Measurement, 5, 363-390.

[60] McFadden, D. (1981) "Econometric Models of Probabilistic Choice," in Structural Analysis of Discrete Data, ed. by C. F. Manski and D. McFadden. Cambridge, Massachusetts: M.I.T. Press.

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

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

[63] Olley, S. and Pakes, A; (1996). *The Dynamics of Productivity in the Telecommunications equipment Industry*. Econometrica 64 (6), 1363-1298.

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

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

[66] Owen, A.B. (1990). *'Empirical likelihood confidence regions'* The Annals of Statis- tics 18, 90-120.

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

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

[69] Orley Ashenfelter (1984). *Macroeconomic Analysis and Microeconomic Analysis of Labor Supply*. Princeton University.

[70] Osborne, Martin (2004) *An Introduction of Game Theory*. Oxford University Press.

[71] Pesendorfer, M., and Schmidt-Dengler, P. (2003): *Identification and Estimation of Dynamic Games*, London School of Economics Working Paper.

[72] Pagan, A and Ullah, A. (1999) *Nonparametric Econometrics*. Cambridge University Press.

[73] Pakes, Ariel, Michel Ostrovsky, and Steve Berry (2004). *Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Samples)*. NBER: Working Paper 10506.

[74] Perloff, Jeffrey, Karp, Larry, and Amos Golan (2007). Estimating Market Power and Strategies. Cambridge University Press.

[75] Pitt, M. and L. Lee, *The Measurement and Sources of Technical Inefficiency in Indonesian Weaving Industry*, Journal of Development Economics, 9, (1981), pp. 43-64.

[76] 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.

[77] 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.

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

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

[80] Rust, J. (1994). *Structural estimation of Markov decision processes.* In R. Engle and D. Mc- Fadden (Eds.), Handbook of Econometrics, Volume, pp. 3081.3143. New York: North- Holland.

[81] Schmith, P. (1981). *Constrains on the Parameters in Simultaneous Tobit and Probit Models.* In C. Manski and D. MacFadden (eds.) *Structural Analysis of Discrete Data with Econometric Applications* (Cambridge, MA: MIT Press) 422-434.

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

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

[84] Tom Domencich and Daniel L. McFadden (1975) *Urban Travel Demand: A Behavioral Analysis* North-Holland Publishing Co., 1975. Reprinted 1996.

[85] Wooldrige M. Jeffrey, 2002. *Econometric Analysis of Cross Section and Panel Data*, MIT press. London.

[86] World Bank, 2002. *Mexico Export Dynamics and Productivity, Analyisis of Mexican Manufacturing in the 1990s.*