The London School of Economics and Political Science

Essays on Industrial Organization

Daniel Silva Junior

February 13, 2014

Thesis submitted to the Department of Economics of the London School of Economics for the degree in Doctor of Philosophy.

London, December 2013

For my mother Maria Rita Silva (in memoriam).

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

I declare that my thesis consists of 36,209 words.

Statement of conjoint work

I confirm that Chapter 2 was jointly co-authored with Sorawoot Srisuma and Fabio A. Miessi Sanches. I contribute together with Fabio A. Miessi Sanches with the main idea. The general directions of the paper were taken jointly with the other two authors. I also contribute with the Monte Carlo simulations. In total, I believe I contribute with one third of this work.

I confirm that Chapter 3 was jointly co-authored with Fabio A. Miessi Sanches. I contribute together with Fabio A. Miessi Sanches with the main idea. The general directions of the paper were taken jointly with the other author. I also contribute with the programming of the structural estimation and part of the writing. In total, I believe I contribute with 50% of this work.

Acknowledgments

I would like to thank Professor Martin Pesendorfer for the support throughout the PhD.

Financial support during my PhD from CNPq (Brazilian Ministry of Science and Technology) is gratefully acknowledged.

I would like to thank my Co-authors Fabio Miessi and Tang Srisuma for proving a essential contribution for this thesis. I also benefited from discussions with Francisco Costa, Dimitri Szerman, Mathew Gentry, Joao Paulo Pessoa, Pasquale Schiraldi, and Jason Garred.

I am indebted to a number of friends that have made my PhD a more pleasant experience: Sofia Galligani, Sutanuka Roy, Carlijn Bogaardt, Ruben Aureliano, Miguel Jose Clavijo, Lucian Todoran, Maria Nicolaou, Valando Nicolaou, Debora Braga Vieira, Daniel Wang, Julianey Cristiny, and Alan Nader. I specially grateful to Stella Hadjiantoni for all the support during the last years of my PhD.

Finally, I would like to thank my family, specially my sisters Andreia Maria Silva and Jussara Silva, for all the support given to me during all these years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis is composed by three essays and provides empirical and methodological contributions to the Industrial Organization literature. The first essay (chapter 2) analyzes the welfare impacts of the Brazilian Biodiesel law. The second essay (chapter 3) develops an alternative estimator for dynamic games. The third essay (chapter 4) applies the methodology developed in chapter 3 to build a dynamic oligopoly model for Brazilian banking industry. Chapter 5 concludes the thesis and provides some directions for future work.

Chapter 1 analyzes market effects of Brazilian biodiesel regulations. Biodiesel was introduced in Brazil in 2005, mixed with mineral diesel to produce the BX blend (X stands for the percentage of biodiesel). Even in small quantities, the percentage of biodiesel has a positive impact on final price of BX because the production costs of biodiesel are higher than those of mineral diesel. In order to analyze the welfare consequences of this price increase, I use a static partial equilibrium framework. The results show that the current proportion of biodiesel in the diesel mixture (5%) increases consumers' price by 1.7% and decreases the consumption by 1.5% compared to the scenario without biodiesel. Also, an increase in the biodiesel percentage to 10% would raise the price by 3.5% and reduce the consumption by 3%.

Chapter 3 provides an alternative estimator for dynamic games. Estimation of dynamic games is a numerically challenging task, in chapter 3 we propose an alternative class of asymptotic least squares estimators to Pesendorfer and Schmidt-Dengler's (2008). The estimator we propose is based on the equilibrium condition of the game when represented in the space of payoffs - in contrast with Pesendorfer and Schmidt-Dengler's (2008) that work in the probability space. Our estimator reduces significantly the computational burden.

This reduction is specially significant under the linear-in-parameter specification where our estimator has an OLS/GLS closed form that does not require any optimization. Also, we show that our estimator is asymptotically equivalent to Pesenrorfer and Schmidt-Dengler's (2008). This implies that there is no theoretical cost of using our estimator. Monte Carlo estimations show that our estimator has good small sample properties and provides significant reduction in the computational time when compared to Pesenrorfer and Schmidt-Dengler's (2008) estimator.

Chapter 4 applies the methodology developed in chapter 3. We estimate a dynamic oligopoly model for the Brazilian banking industry. The results are used to build counterfactuals to examine the effects of the privatization of public banks on the number of bank branches in small municipalities. We find that public banks are not strategic and their presence generate positive spillovers on the private banks' profits. The model however, is not able to disentangle the nature of this spillover. Also, the counterfactual shows that the number of branches operating in small markets would drop in a privatization scenario.

Chapter 5 is the conclusion, I discuss the limitations of the current work and provide some directions for future research.

Chapter 2

Impacts of Biodiesel on the Brazilian fuel market

The paper is published in Energy Economics 2013 (36) 666-675

Abstract

This paper investigates market effects of the Brazilian biodiesel law, which made the use of biodiesel, blended with petroleum diesel, mandatory in Brazil. The study estimates the demand curve for diesel fuel (biodiesel and petroleum diesel) and the industry supply curve of biodiesel. These two pieces of information have been used in a static analysis to draw scenarios with different biodiesel mandates. The results show that the current proportion of biodiesel in the diesel mixture (5%) increases consumers' price by 1.7% and decreases the consumption by 1.5% compared to the scenario without biodiesel. Also, an increase in the biodiesel percentage to 10% would raise the price by 3.5% and reduce the consumption by 3%.

2.1 Introduction

Recent concerns about the environment, high fossil fuels prices, and energy security led to the creation of biodiesel programs in several countries (Althoff, Ehmke,and Gray (2003),Lamers, McCormick, and Hilbert (2008), Ayhan (2007)). In Brazil, biodiesel was introduced in 2005

(Brazil (2005)), mixed with mineral diesel to produce the BX blend, where X stands for the percentage of biodiesel.

During a transition period, from 2005 to 2007, a 2% addition of biodiesel to the mineral diesel (B2) was optional. After that, a certain percentage of biodiesel was required in all the diesel fuel[1] sold in Brazil. The biodiesel mandate has increased over time (Brazil (2005)). In the first six months of 2008 the commercial diesel fuel had to contain 2% of biodiesel (B2). The biodiesel proportion was required to increase to 3% (B3) on 1 July 2008, to 4% (B4) one year later, and to 5% (B5) since January 2010 (three years before the initial goal, MME (2005)).

Brazilian federal government gave ecological, economic and social reasons to introduce the biodiesel mandate (MME (2005)). From the environmental perspective, the biodiesel is supposed to have smaller impact on greenhouse gas emissions since it is a renewable fuel produced mainly from vegetable oils (Ayhan (2007)). From the economic point of view, the fuel was expected to diminish petroleum diesel importation. In 2005 about 6% of the petroleum sold in the Brazilian was imported. Also, Brazil could become an exporter of biodiesel since other countries are adopting similar programs. Finally, the biodiesel production could be an instrument to reduce regional inequality through income and employment generated by the biodiesel production chain. Fiscal benefits were implemented for all biodiesel producers who use raw materials from small farmers in poor regions.

However, biodiesel adoption may have negative impacts. First, it could increase the emissions of greenhouse gases if the production causes deforestation (Fargione et al. (2008)). Secondly, biofuel demand increases feedstock prices (Pfuderer, Davies, and Mitchell) and may cause land use change. For example, Almeida, Bomtempo, and Silva (2008) show some evidence of substition of traditonal crops (such as oranges) to sugarcane in São Paulo state. Also, biodiesel may cause loss of welfare as its current production costs are higher than petroleum diesel (Althoff, Ehmke,and Gray (2003)). This is a non-exhaustive list as other problems might be caused by biodiesel use.

Since biodiesel adoption generates benefits and problems it is necessary to produce studies quantifying these effects. The need is even more evident when one analyzes the economic importance of the diesel. The fuel is the base of Brazilian transportation system, used in trucks, boats, buses, tractors and some power plants – the last mainly in North (Amazon)

[1] The term “diesel fuel” in this article refers to the biodiesel-diesel mixture. The terms “mineral diesel”, “petroleum diesel”, and “oil diesel” are used as synonyms.

region. Even in small proportion, biodiesel may have significant impacts on the Brazilian Economy.

The present paper contributes to the biodiesel literature by providing a measure of the welfare impact of biodiesel adoption. The empirical strategy consists in estimating the demand for diesel fuel and the costs of the biodiesel industry. With these two pieces of information it is possible to draw scenarios changing the biodiesel mandate. Then, the scenarios can be used to evaluate the effects of the biofuel plan and to predict the impact of future changes.

The economic impacts quantified in this paper are based on traditional partial equilibrium analysis: a change in the biodiesel policy changes the market equilibrium (price and quantity), which implies a change in the consumer and producer surplus [2]. De Gorter and Just (2009) provide a general framework for the use of partial equilibrium analysis in the biofuel context.

The demand estimate is based on Brazilian States' monthly panel data. The period analyzed spans from January 2003 to December 2009. The data include diesel prices, amount of diesel sold (pure diesel before 2008 and BX mixture onwards), the fleet of heavy vehicles, and ICMS (tax on trade of products and services) as a proxy for economic activity.

The biodiesel industry supply curve was approximated with data from the last (17^{th}) biodiesel auction. This information is combined with the average price of petroleum diesel in the refineries, and with wholesale and retail prices of the biodiesel mixture to simulate the impacts on price and quantity of BX sold caused by changes in the biodiesel mandate.

The results obtained show that the current proportion of biodiesel in the diesel mixture (B5) increases consumers' price by 1.7% and decreases the consumption by 1.5% when compared to the environment without biodiesel. An increase in the biodiesel percentage to 10% (B10) would raise the price by 3.5% and reduce the consumption by 3%. These all add up to considerable welfare loss to consumers, retailers, and wholesalers.

The paper builds on a static approximation of a dynamic process: industry capacity and marginal costs are held constant over the simulations, and we ignore entry cost, adjustment costs and strategic price behavior.

The biodiesel industry problem, however, is intrinsically dynamic. Firms have to decide whether to enter the market or not. Once a firm has entered, it has to set the capacity, the

[2] A comprehensive introduction to welfare analysis can be found in Mas-Colell, Whinston,and Green (1995)

technology, and the location and, in each period of time, it decides to continue or to exit the market. Firms decisions will affect the market competition and consequently the prices. Therefore, a natural extension of this paper is a dynamic oligopoly model in the tradition of Ericson and Pakes (1995), and Maskin and Tirole (1988). The model must deal with intrinsic heterogeneity in the industry regarding capacity, location and technology employed.

The paper is organized as follows. The next section describes the biodiesel market and shows the dataset used in the paper. Section 2.2 estimates the demand curve for diesel fuel. Section 2.4 recovers the costs and estimates the supply curve for the biodiesel industry. Section 2.5 presents the results of the simulations. Finally, section 2.6 concludes the paper, provides some directions for future work, and discusses some limitations of the approach.

2.2 The Biodiesel Market

Brazil has a long experience in the use of biofuels. During the 1970's the Proalcohol program (Brazil (1975)) developed bioethanol as a substitute to gasoline in automobiles. Even though the legislation has experienced several changes over the last four decades, ethanol is still a very important part of Brazil's energy matrix (ANP (2010)), and in 2006 ethanol represented 17% of Brazilian fuel supply (Almeida, Bomtempo, and Silva (2008)).

Biodiesel, on the other hand, was adopted later in Brazil. Silva César and Batalha (2010) show that a first attempt to implement biodiesel production was made in 1980. However, it was abandoned in 1986 due to reduction in the petroleum barrel price. During the 2000s new concerns about renewable energy led to the creation of the National Program for the Production and Use of Biodiesel (PNPB) (MME (2005)). The main result of the program was law n°11.097/2005 (Brazil (2005)), which made the use of biodiesel mandatory from 2008.

It is worth noting that ethanol and biodiesel are not market competitors. Ethanol is a substitute to gasoline, used basically in automobiles. Biodiesel is a substitute to oil diesel, used mainly in trucks and buses. Since 1976, the use of diesel engines in automobiles has been forbidden by law (Brazil (1976)). Therefore, ethanol belongs to a different market and is not analyzed in the present paper.

The recent introduction of biodiesel generated a growing literature about the topic. Silva César and Batalha (2010) summarize the history of the biodiesel in Brazil. Ayhan (2007) discusses some benefits of biodiesel and governmental policies regarding biodiesel. Barbosa (2011) and Pfuderer, Davies, and Mitchell (2009) look at the impacts of biofuels in the feedstock market. Fargione et al. (2008) analyse the impact of biofuels in the greenhouse gas emissions.

From the economic perspective, it is possible to use partial equilibrium analysis, a useful economic tool, to assess market outcomes of the use of biodiesel. De Gorter and Just (2009) propose a general framework to analyze the impact of different biofuel mandates alongside taxes. The authors show how different mandates and tax schemes impact prices, and use data to recover supply elasticities from gasoline and ethanol. Similarly, Althoffff, Ehmke, and Gray (2003) use partial equilibrium analysis to quantify the losses to the Indiana economy caused by a 2% biodiesel mandate. Their estimative shows a total cost ranging from $ 15.2 to 17.2 million.

In this tradition, the present paper contributes to the literature by using partial equilibrium analysis to quantify the impact of the biodiesel mandate on the market equilibrium outcomes. Due to the rich dataset employed in the paper it is possible to show how equilibrium prices and quantities are changed according to the biodiesel mandate, given that other factors remain constant (see section 2.5). This also provides some quantitative measures of biodiesel's impact on welfare.

2.2.1 Commercialization

The Brazilian oil and biofuel market is regulated by ANP[3] (Brazil (1997)). For biodiesel specifically, the agency is responsible for determining biodiesel standards, inspecting the market (to assure that the correct biodiesel mandate is sold), and for collecting data. Also, ANP provides licences to construct new biodiesel plants, to change the capacity of existing ones, to produce, and to commercialize the biofuel.

ANP also plays a direct role in biodiesel commercialization. Wholesalers are responsible for mixing biodiesel and oil diesel (ANP (2010)). However, they are not allowed to negotiate directly with the producers. Instead, they have to buy biodiesel through actions organized by ANP.

ANP determines the amount of biodiesel that must be sold and the auction rules. The biodiesel producers are the bidders. They bid a mix of price and quantity according to the

[3]National Agency of Petroleum, Natural Gas and Biofuels

specific rules of the auction. The winners are those bidders with the lowest prices. The buyers are oil refineries[4]. Each refinery is assigned a percentage of the total volume of biodiesel by ANP. After the auction the oil refineries resell the biodiesel to the wholesalers (ANP (2010)).

Seventeen auctions were performed between 2005 and the first quarter of 2010 (ANP (2010)). From the 12^{th} auction onwards ANP divided each auction in two. The first part of the split auctions were restricted, only the bidders that bought raw material from small farmers were allowed to participate. In the second part, all registered producers could participate as bidders (see Silva César and Batalha (2010) for details). Where necessary, the split auctions have a different notation through the paper: after the ANP number they have a (i) symbol, where $i = 1$ means that the auction is restricted and $i = 2$ means that the auction is non restricted. Also, they are considered different auctions because they have a separated dynamics.

Table 2.1 summarizes the results of the biodiesel auctions. It can be seen that the number of bidders and the volume auctioned increased over time, reflecting the increase in the biodiesel mandate. The prices, on the other hand, did not follow this pattern. Both the ceiling and the average price increased from the 6^{th} to the 12^{th} auctions and then the prices returned to the initial levels.

Empirical analysis of the biodiesel auctions is virtually impossible due to their peculiarities [5]. Firstly, there were drastic changes in the auction format. The first eight auctions were electronic auctions; ANP implemented live auctions for the following eight auctions and returned to electronic format for the last one. Secondly, the bidding rules, the regularity of the auctions, the delivery schedule, and the guarantee of producers also changed over time. Finally, the number of auctions is not large enough for empirical analysis.

However, even without an accurate analysis, the results of the auctions are used in section 2.4 to recover information regarding the supply curve of the biodiesel industry.

2.2.2 Market Data

ANP collects fuel market data on a monthly basis. The information includes the amount of each fuel sold by state, the total quantity of each fuel produced by oil refinery, the total

[4] Technically, the buyers in the auctions are all the producers and importers with market share above 1%. In practise, the oil refineries of two companies (Petrobras and Refap) fulfill these conditions (ANP (2010)).

[5] For empirical estimation of auctions see Guerre, Perrigne, and Vuong (2000), Laffont, Ossard, and Voung (1995), and Donald and Paarsch (1993)

Table 2.1: Results of Biodiesel Auctions

Auction	Date	Price (R$ Jan 2003/m3)		Volume Auctioned (m3)	N°of Bidders
		Celing	Average		
1	23/11/2005	1624.80	1611.97	70000	8
2	30/03/2006	1593.16	1552.78	170000	12
3	11/07/2006	1588.07	1462.13	50000	6
4	12/07/2006	1587.79	1456.19	550000	25
5	13/02/2007	1555.97	1521.36	45000	7
6	13/11/2007	1907.07	1483.61	304000	26
7	14/11/2007	1907.07	1480.52	76000	30
8	10/04/2008	2165.74	2078.10	259000	24
9	11/04/2008	2165.74	2074.01	66000	20
10	14/08/2008	1977.55	1965.95	264000	21
11	15/08/2008	1977.55	1969.77	66000	20
12(1)	24/11/2008	1794.61	1784.83	264000	23
12(2)	24/11/2008	1794.61	1787.96	66000	23
13(1)	27/02/2009	1737.39	1636.30	252000	27
13(2)	27/02/2009	1737.39	1387.99	63000	32
14(1)	25/05/2009	1712.05	1673.58	368000	27
14(2)	25/05/2009	1712.05	1680.81	92000	32
15(1)	27/08/2009	1657.58	1631.36	368000	27
15(2)	27/08/2009	1657.58	1639.82	92000	32
16(1)	17/11/2009	1685.85	1670.46	460000	29
16(2)	17/11/2009	1685.85	1663.75	115000	34
17(1)	12/02/2010	1614.04	1572.47	419900	29
17(2)	12/02/2010	1614.04	1556.73	106000	43

Notes: The data is from National Agency for Oil and Biofuels ANP (2010).

volume of fuel imported, average fuel prices by state (price collected over a significant sample of gas stations) and the average price charged by producers and importers of petroleum products at regional level[6].

The data on quantity of fuel sold start in January 2000. All wholesale fuel distributors have to report to ANP the amount of fuel sold per month in each state. The data, therefore, cover all Brazilian territory and include all types of liquid fuels used in the market.

Data on prices started being collected by ANP in July 2001 and the number of cities and gas stations consulted has increased over time. In July 2001 the research covered 411 cities, the total increased to 555 municipalities in May 2004.

Data on the total fleet of buses (including those used in public transportation), tractors and trucks by state from January 2003 to December 2009 were also collected. This information is from Denatran (National Department of Traffic) and it is available at the state level. Besides, as a measure of monthly economic activity level, information was gathered on ICMS (Tax on Trade of Products and Services) for the same period. All prices used in the paper are adjusted to January 2003 constant reais (Brazilian currency).

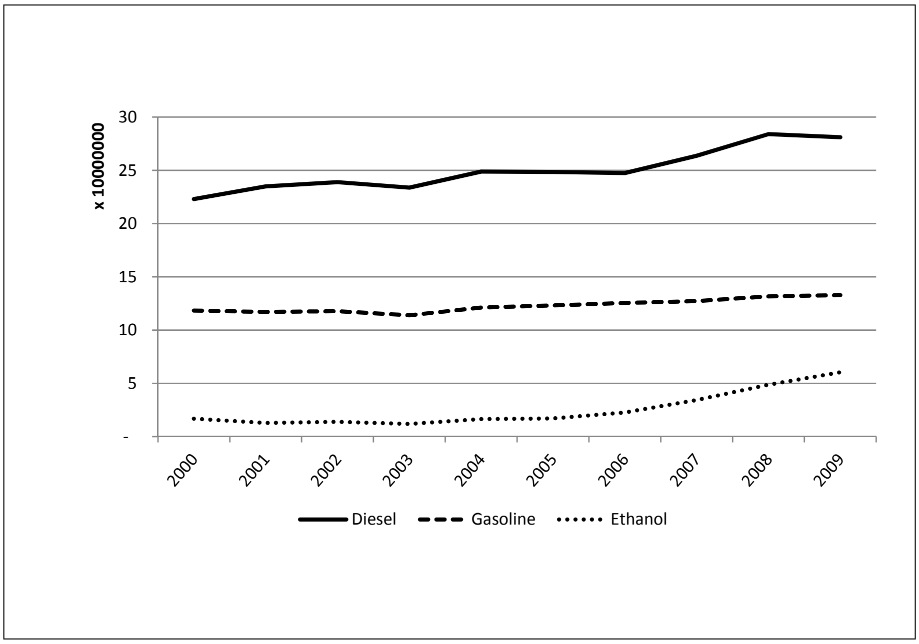
To serve as instruments for demand estimation I obtained data on the total value of petroleum imports and on the average wage of new employees in the wholesale fuel distribution and in the fuel retail industries. The value of imports is from ANP and is available with a monthly frequency at the national level. The wages are from the Brazilian Ministry of Labour and Employment; this data have monthly frequency and it is available at the state level.

Figure 2.1 shows the consumption of the three main liquid fuels used in land transport in Barrel of Oil Equivalent (BOE) from 2000 to 2009. As said before, gasoline and ethanol are not substitutes to diesel in the Brazilian market. However, the graph illustrates the importance of diesel fuel. Diesel is consumed more than gasoline and ethanol put together. Nevertheless, ethanol has the highest increase in consumption (250% in the period), followed by the diesel (26%) and gasoline (12%).

Table 2.2 shows the consumption of diesel per Brazilian Region in 2007. In terms of absolute consumption, the Southeast uses most of the diesel, approximately 43% of the total consumption. Among the states, São Paulo, in the Southeast region, has the highest consumption, about 22% of the national consumption. However, considering consumption

[6]Each region is formed by groups of states with geographical, historical and economic similarities.

Figure 2.1: Gasoline, Ethanol and Diesel Consumption in BOE



Source: Elaborated based on data from the National Agency for Oil and Biofuels ANP (2010).

per capita, the Central-West region has the highest use of diesel followed by the South region. The Southeast region has the second lowest per capita consumption. Table 2.2 provides evidence that the regional effects are important for the diesel fuel demand.

Table 2.2: Diesel Consumption by Region

	Consumption (1000000 BOE)	Consumption per capita
North	3.87	0.26
Northeast	6.23	0.12
Central-West	4.67	0.35
Southeast	18.10	0.23
South	8.68	0.32

Notes: The data are from National Agency for Oil and Biofuels ANP (2010).

Figure 2.2 shows the diesel fuel price dynamic in Brazil. The left hand side of the figure shows the average price charged by retailers, the minimum and the maximum price found in the survey and the average retail margin (the difference between the price charged by the gas stations and the wholesale price). The right hand side of the figure shows the standard deviation of the prices.

Figure 2.2: Price (R$ January 2003)

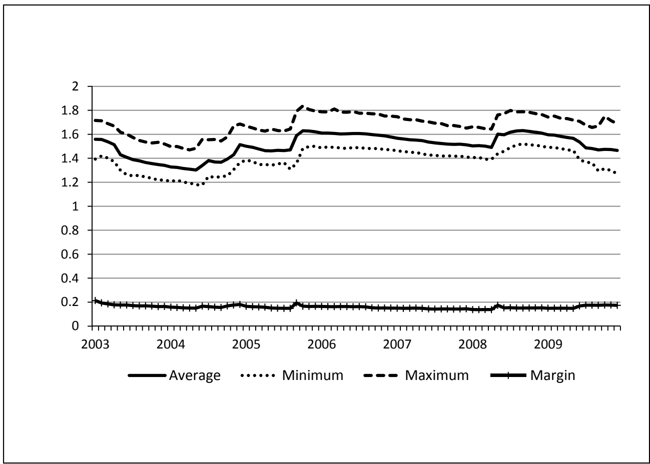
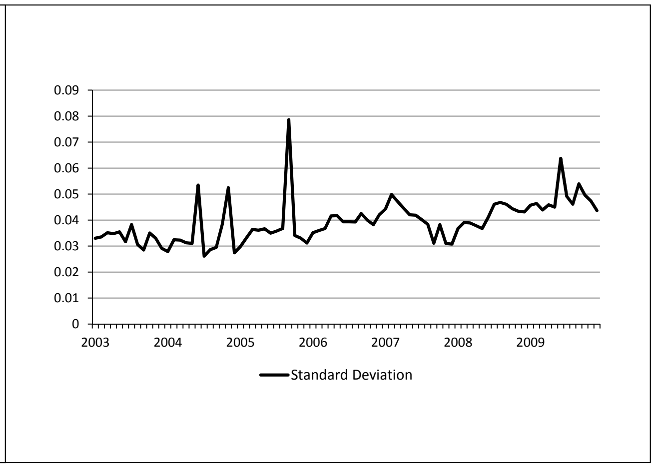


Figure 2.3: Standard Deviation



Source: Elaborated based on data from the National Agency for Oil and Biofuels ANP (2010).

One can see that the movements of the average, minimum and maximum price are similar with a reasonably stable difference among them. The variance of the price reinforces this feature. It is small compared to the average, less than 0.05 most of the time, with a slight increase at the end of the period.

The average retail margin is stable during the whole analyzed period, indicating that the gas stations are able to transfer the positive shocks on costs to consumers. Conversely they do not retain a negative shock on prices. The figure indicates that the gas stations are charging a fixed markup over the marginal cost.

If one does not consider the retail margins of the other fuels, the result on stable margins differs from those found by Hosken, McMillan, and Taylor (2008). The authors discovered a substantial variability in the retail margins. This difference, however, may be explained by data aggregation and might not reflect the individual behavior of the retailers.

Table 2.3 shows some basic statistics regarding the biodiesel industry. All the indicators have grown over the period. The number of plants in 2009 is more than nine times the total number of factories in 2005. The production experienced an even higher increase, more than 2000 times the initial quantity. In 2005 the production by the Brazilian biodiesel industry was only 756 m^3 and it rose to 1.6 million m^3 in 2009.

The capacity, on the other hand, had a much slower increase when compared to the production. The capacity was around 85,000 m^3 in 2005 and increased to about 460,000 m^3 in 2009 (almost 55 times the initial capacity). The capacity utilization, however, is still low,

Table 2.3: Industry Summary Statistics

	2005	2006	2007	2008	2009
Production (1000 m^3)	0.7	69.0	404.3	1,167.1	1,607.8
Number of Plants	7	18	45	62	64
Capacity (1000 m^3)	84.7	638	2,475.1	3,315.3	4,629.8
Capacity Utilization (%)	0.9%	10.8%	16.3%	35.2%	34.7%

Source: Elaborated based on data from the National Agency for Oil and Biofuels ANP (2010).

around 35% in 2009.

This growth in the indicators was expected as the percentage of biodiesel in the BX also increased. However, the low percentage of the capacity utilization raises questions about the firms' strategy and about the future of the industry market structure. The capacity production of biodiesel is enough to attend levels of BX higher than 5%. Therefore, in the absence of new increases in the biodiesel mandate, some firms are expected to exit the industry.

2.3 Demand Estimation

The demand for diesel is a result of a number of different maximization processes. Diesel is used as input in several industries: agriculture, land transportation (freight and passengers), ship transportation and energy generation. In addition to industrial use, diesel can also be utilized by domestic consumers for idiosyncratic reasons (for example, small boats for recreational fishing). In this sense, the estimation of demand for diesel cannot rely on a structural model.

Following the considerations above, the demand for diesel can be seen as a function of variables such as economic activities and fleet level.

$$lnQ_{it} = f\left(p_{it}, X_{it}\right) + \varepsilon_{it} \tag{2.3.1}$$

Where lnQ_{it} represents the natural logarithm of the quantity (m^3) of diesel sold in state i in period t, p_{it} is the average price charged by gas stations in state i during the period t. The vector X_{it} is composed by covariates that influence demand. It includes the logarithm of ICMS as a measure of economic activity level, the logarithm of the total fleet (the sum of the buses, tractors and trucks by state) to capture the importance of diesel in transportation, and state dummies as regional characteristics may affect the demand for diesel fuel. To mitigate endogeneity problems with the variable ICMS, it does not include fuel taxes. ε_{it} is the error term. $f(.)$ is a demand function.

A similar approach to estimate gasoline demand has been used in many studies as summarized by Basso and Oum (2007). The traditional approach takes the demand for gasoline as a function of price, income and controls. Also, most of the specifications rely on log-linear forms and Greene (1982) and Dahl and Sterner (1991) support the selection of the log-linear form. Based on that, the present paper also estimates a log-linear specification for the demand function. Equation 2.3.1 becomes:

$$lnQ_{it} = \alpha_0 + \alpha_1 ln(p_{it}) + X_{it}\beta' + \tau_i + \lambda_t + \upsilon_{it} \tag{2.3.2}$$

ln represents the natural logarithm and the error $\varepsilon_{i,t}$ is decomposed into three terms: an individual specific term τ_i, a time specific term, λ_t, and an individual time specific term ,$\upsilon_{i,t}$.

In order to assure the consistency of demand parameters estimated, one assumption is necessary:

D1: *Consumers are only interested in the amount of energy the fuel produces.*

Assumption **D1** says that consumers see any biodiesel mandate as the same product [7]. This assumption assures that the coefficients of equation 2.3.2 are stable, that they do not change according to the diesel mandate.

Table 2.4 presents the results of the demand estimation. The first column of the table is the result of the OLS regression with the use of state dummies, which is equivalent to the fixed effects estimator [8]. The coefficient of the price is negative, as expected. The value is approximately 0.6 indicating that an increase of one percent in the real price of diesel implies a reduction of 0.6% in the consumption of diesel. The variable fleet has an unexpected negative signal: an increase in the number of heavy vehicles decreases the use of biodiesel. However, the coefficient is statistically insignificant. On the other hand, the log of ICMS has the expected signal. A one percent increase in the total tax collected causes an increase of 0.03% in the total consumption of diesel. Since ICMS is charged on products

[7] Tests conducted by the Ministry of Science and Technology indicated that there are no loss of efficiency in diesel engines due the use of any BX blend up to B5 (ANP (2010))

[8] See Greene (2003) and Wooldridge (2002) for details about panel methods

and services effectively traded, the results show a positive relation between economic activity and diesel consumption.

The second column in table 2.4 shows the results of the random effects estimation. One can see that the price and economic activity effects are stronger under the random effects hypothesis. Furthermore, the fleet has the expected sign and it is significant at 5%. However, the Hausman test rejects the hypothesis of no systematic differences between random and fixed effects.

Instrumental variables (IV) are used to control the endogeneity problem caused by simultaneous equations. The instruments are supply-side cost shifters: the log of the wholesale average price, the log of the wage of new employees in the fuel distribution industry (retail and wholesale), and the log of average import price per m^3 of petroleum. All instruments affect the supply of diesel fuel, as they affect the marginal cost of the industry. However, they have no effects on the demand side. In other words, the instruments do not alter the consumers' decisions. Therefore, the instruments can be considered exogenous. The regression of the retail price on the instruments (table 2.5) has an F-statistic equal to 9496, considerably higher than the 10 or 20 value pointed by Stock, Wright, and Yogo (2002) to rule out weak instruments.

The results for the IV with state dummies can be found in the third column of table 2.4. The price elasticity is considerably higher when compared to OLS regression: the value of the parameter is now about -0.9, 50% higher than the estimation without instruments. The log of fleet continues to be insignificant and with the wrong sign. The log of ICMS is almost the same when compared with the OLS regression.

Finally, the fourth column in table 2.4 presents the random effects estimation of the model using instrumental variables. The price elasticity is around one and the log of fleet is significant and with the right sign. The log of ICMS is close to the value obtained in the random effects estimation with no instruments and higher than the fixed effects estimations. This difference between the estimators with instruments is strongly significant according to the Hausman test (chi-square value of 407.12).

One can draw two conclusions based on the diesel fuel demand estimation. First, the regional effects are important. The Hausman test strongly rejects equality among the fixed and the random effects estimators [9]. Second, the use of instrumental variables changed the

[9] The random effects estimator is consistent and efficient under the hypothesis of independence of the individual characteristics. The fixed effects estimator does not need this assumption to achieve consistence.

Table 2.4: Demand Estimation

	OLS	RE	IV	RE-IV
Log Price	-0.5599*	-0.6398*	-0.8594*	-1.0573*
	(0.1701)	(0.1760)	(0.1998)	(0.1998)
Log Fleet	-0.0123	0.0535*	-0.0109	0.0521*
	(0.0171)	(0.0171)	0.0170)	(0.017)
Log ICMS	0.0394*	0.0730*	0.0417*	0.0748*
	(0.0126)	(0.0128)	(0.0125)	(0.0127)
State Dummies	Yes	No	Yes	No
Time Dummies	Yes	Yes	Yes	Yes
Number of Observations	2259	2259	2257	2257
R-Squared	0.9864	0.6723	0.9866	0.6634
Hausman	402.98		407.12	

Notes: The dependent Variable is the log of the total m^3 of diesel. Standard errors in parentheses. *Significant at 5%. Instruments are the log of the wholesale price, the log of the import expenditure of petroleum and the log of the wage of new employees in the fuel distributor industry.

Table 2.5: Regression of the Instruments on the Logarithm of the Price

	OLS
Ln Wholesale Price	0.9111*
	(0.0063)
Ln New wages	0.0020
	(0.0004)
Ln Oil Import Price	-0.0073*
	(0.0027)
F-Statistic	9496

Notes: The dependent variable is logged diesel price. Standard errors in parentheses. *Significant at 5%.

results considerably. The OLS seems to underestimate the price elasticity.

Based on the conclusions above, the coefficients obtained in the fixed effects IV estimation are used to construct the simulations in section 6.

2.4 The Production Side

Data on the last two auctions (17(1) and 17(2)) was used to approximate the industry supply curve. The are two reasons for this choice. First, these two auctions are electronic and therefore have better information regarding the bids; it is possible to access all the bids of every firm. Second, the paper relies on a static exercise and to include the past auctions I should make considerations regarding the capacity adjustments cost and the entry costs of the firms.

Since the cost structure of the firms is not observed, some assumptions are required to recover the supply curve. First, I assume a strategic interaction among the firms and a feature of the non-observed cost structure:

P1: *Firms are in perfect competition.*

P2: *Marginal cost is constant and firms can produce up to 100% of their capacity.*

Assumption **P1** is based on the low capacity utilization in 2009, around 36%. The firms are obliged to enter into a fierce competition in order to sell their production. The first part of assumption **P2**, constant marginal costs, is a standard assumption in both theoretical[10] and empirical[11] economic literature. The second part of assumption **P2** refers to the fact that firms do not waste money building a capacity that they will not use.

Assumptions **P1** and **P2** are, however, not enough to characterize the supply curve as some firms did not enter in the the last two auctions. To overcome this difficulty, the firms are divided in four groups, according to their participation in the auction, and specific assumptions are made for each group.

Therefore the fixed effect is a more robust estimator. See Wooldridge (2002) for use of the Hausman test in the panel context.

[10]For example, Sutton (1991), Krugman (1979), Dixit and Stiglitz (1977) and Tirole (1988).

[11]For example, Ryan (2012), Nevo (2001), and Jia (2006).

2.4.1 Group 1

The first group is composed of the firms that had license to produce but not to commercialize. They could not have participated in the auctions. For this group the following assumption is made:

P3: *The marginal cost is equal to the auction ceiling price. Therefore, firms in group 1 can offer any amount of biodiesel between zero and their full capacity at the ceiling price.*

Formally:

$$s_i(b) = \begin{cases} q_i = 0 & \text{if } b < b_c \\ 0 \leq q_i \leq C_i & \text{if } b = b_c \\ q_i = C_i & \text{if } b > b_c \end{cases} \tag{2.4.1}$$

Where $s_i(b)$ is firm i's supply function, C_i is firm i's capacity, q_i is firm i's quantity supplied, b_c is the auction ceiling price, and b is a given price.

Equation 2.4.1 may underestimate or overestimate the supply function of the plants in this group as no information about the price behavior is known. However, this problem is minimized as a small fraction of firms belong to this group (see appendix).

2.4.2 Group 2

The second group consists of firms that could have entered the auction (they have the commercialization license) but decided not to enter. The assumption made for this group is exactly equal to **P3**:

P4: *The marginal cost is equal to the auction ceiling price. Therefore, firms in group 2 can offer any amount of biodiesel between zero and their full capacity at the ceiling price.*

Formally:

$$s_i(b) = \begin{cases} q_i = 0 & \text{if } b < b_c \\ 0 \leq q_i \leq C_i & \text{if } b = b_c \\ q_i = C_i & \text{if } b > b_c \end{cases} \tag{2.4.2}$$

Where $s_i(b)$ is firm i's supply function, C_i is firm i's capacity, q_i is firm i's quantity supplied, p_b is the auction ceiling price, and b is a given price.

For this group the minimum offer price may be higher than the ceiling price. The ceiling price can be seen as a lower bound for the minimum offer price. **P3** implies that the lower bound is actually equal to the minimum offer price for group 2 and the firms did not enter the auction due to entry costs [12].

2.4.3 Group 3

This group incorporates the firms that entered in the auction but did not win. For group 3 the following assumption is made:

P5: *The marginal cost is equal to the firm's lowest bid. Therefore, firms in group 3 can offer any amount of biodiesel between zero and their full capacity at a price equal to their lowest bid.*

Formally:

$$s_i(b) = \begin{cases} q_i = 0 & \text{if } b < b_m^i \\ 0 \leq q_i \leq C_i & \text{if } b = b_m^i \\ q_i = C_i & \text{if } b > b_m^i \end{cases} \tag{2.4.3}$$

Where $s_i(b)$ is firm i's supply function, C_i is firm i's capacity, q_i is firm i's quantity supplied, p_m^i is the firm i's lowest bid, and b is a given price.

Assumption **P3** is a good approximation of the marginal cost as the firms signaled their to provide biodiesel at this price level. Also, assumption **P2** assures that this marginal cost is the same for the entire capacity.

2.4.4 Group 4

This group consists of firms that won the auction. For this group the following assumption is assumed:

P6: *The marginal cost is equal to the firm's lowest winning bid. Therefore, firms in group 4 can offer any amount of biodiesel between zero and their full capacity at a price equal to their lowest winning bid.*

[12] see Li and Zheng (2009) for entry costs in auctions

Formally:

$$s_i(b) = \begin{cases} q_i = 0 & \text{if } b < b_w^i \\ 0 \leq q_i \leq C_i & \text{if } b = b_w^i \\ q_i = C_i & \text{if } b > b_w^i \end{cases} \tag{2.4.4}$$

Where $s_i(b)$ is firm i's supply function, C_i is firm i's capacity, q_i is firm i's quantity supplied, b_w^i is the firm i's lowest winning bid, and b is a given price.

For this group the minimum offer price may be lower than the winning bid. Therefore, their marginal cost might be overestimated.

In the next subsection we combine the assumptions made above to recover the industry supply curve.

2.4.5 Industry Supply

The industry (or market) supply curve is the horizontal sum of firm supply curves (Mas-Colell, Whinston, and Green (1995)). For the biodiesel industry it can be defined as follows:

$$S_i(b) = \sum_{i=1}^{J} s_i(b) \tag{2.4.5}$$

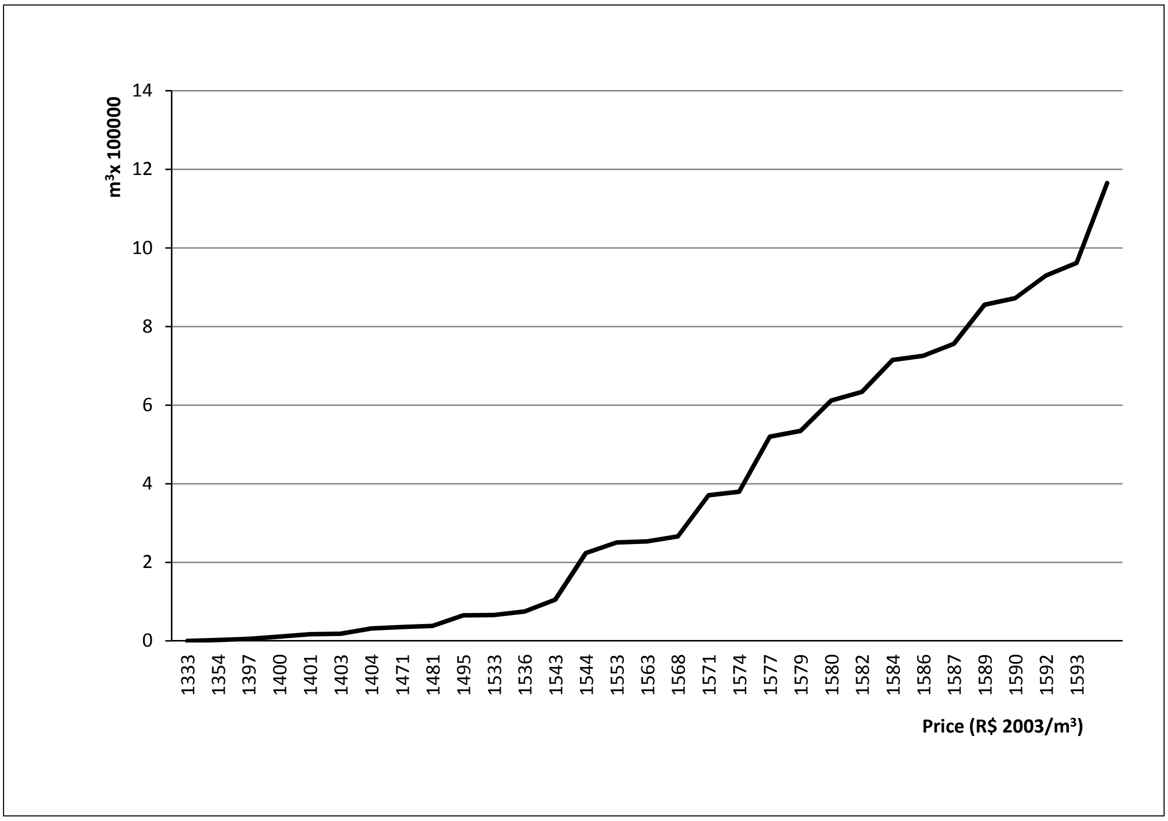
Where $S_i(b)$ is the market supply function, $s_i(b)$ is firm i's supply function, b is a given price, and J is the total number of producers.

Due to the discontinuous characteristics of the firms' supply, the industry supply function does not have a closed form. However, under the assumptions assumed, it is possible to compute the amount of biodiesel offered at any given price. Figure 2.4 shows the biodiesel supply curve.

The supply curve expresses a *ceteris paribus* condition: it shows the relations between the price of a good and its quantity supplied, given that the other factors are constant. Therefore, a change in the these other factors shifts the position of the supply curve. A linear cost reduction for all firms, for example, shifts the demand to the left. For any given price there is a reduction in the quantity offered. The opposite happens if the costs increase.

The factors that affect the biodiesel supply include the raw material price, the opportunity cost of the producer, labor cost, and technology. Biodiesel is produced by a chemical reaction of lipids (vegetable oil or animal fat) with an alcohol (Ayhan (2007)). Therefore, if the price

Figure 2.4: Biodiesel Industry Supply Curve



Source: Elaborated based on data from the National Agency for Oil and Biofuels ANP (2010).

of a feedstock used in biodiesel production increases, the supply curve shifts to the right. Also, if the producer can use the plant to produce oils for non-fuel purposes, a decrease in the price of this alternative option would shift the supply curve to the right. A similar reasoning can be made for all the relevant factors shifting the supply curve.

Even though the factors affecting the supply curve are very important, they are not addressed in this paper for two reasons. First, there is no data available to map the factors to the cost structure. In other words, the quantitative result of these changes cannot be determined. Second, the paper makes a short run partial equilibrium analysis. I analyze how the market outcomes change with a change in the biodiesel mandate given that the other factors are constant (see section 2.5). The hypothesis of all other factors remaining constant is not strong in the short run, as the producer of other goods would take time to adjust prices.

2.5 Welfare Analysis

Biodiesel producers are part of a broader fuel industry which includes wholesalers, retailers, oil refineries and fuel importers. To analyze the impacts of biodiesel mandates on the retail prices it is necessary to consider the interactions among all these agents. More specifically, it is necessary to know how all the agents in the market react to a change in biodiesel mandate.

It is useful to divide the agents in the market according to their position in the diesel fuel supply chain. Firms involved in direct production or importation of the fuel (oil refineries, fuel importers and biodiesel producers) form the upstream part of the supply chain. On the other hand, firms that commercialize the fuel previously produced (wholesalers and retailers) form the downstream part of the supply chain.

The analysis is done backwards. The first step is to see what happens in the downstream part of the market. In other words, how retailers and wholesalers react to a given change in the prices of upstream firms [13]. Figure 2.5 shows the average prices of petroleum diesel charged by retailers, wholesalers, producers and importers in Brazil between January 2003 and December 2007, before the mandatory adoption of biodiesel. The lines of the three prices show a quite similar pattern, also the distance between the retail and the wholesale price is stable over the period. It corroborates the points discussed in section 2.2.2 about the retail markups. The wholesale margin is not so stable; it seemed to increase in the last months of the analyses.

Figure 2.5 provides the base for the following hypothesis:

S1: *(i) Wholesalers and retailers charge a markup (margin) over their acquisition costs; (ii) wholesalers buy petroleum diesel, while biodiesel and retailers buy the BX blend; (iii) the markup at a given point in time is composed of a time invariant part and a time specific error, independent and identically distributed (i.i.d.) with zero mean.*

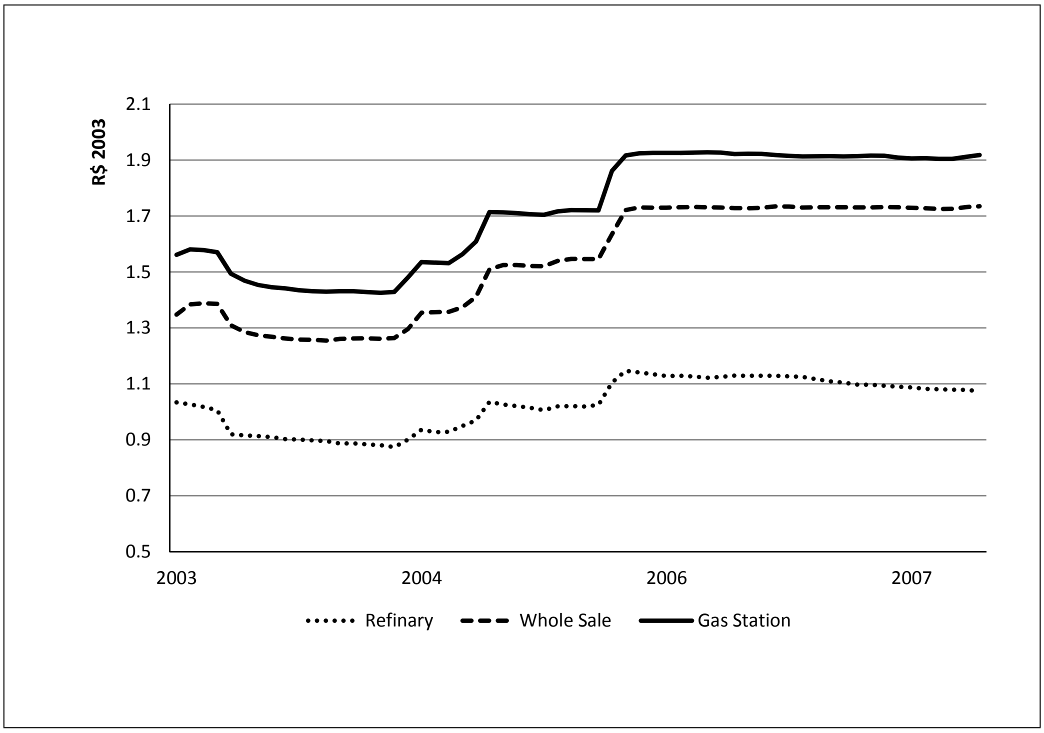
Hypothesis **S1** provides the best response of retailers and wholesalers to any price chosen by firms in the upstream market. It is worth noting that the hypothesis fits the data and simplifies the strategies of the downstream firms.

For the upstream market the following hypothesis is made:

S2: *Oil refineries and diesel importers do not react (change prices) to a change in the biodiesel mandate.*

[13] See Tirole (1988) for a more detailed discussion of vertical restraints.

Figure 2.5: Average Diesel Prices (R$ January 2003)



Source: Elaborated based on data from the National Agency for Oil and Biofuels ANP (2010).

Hypothesis **S2** is reasonable, as the data does not show any change in the oil diesel producers' price caused by the introduction of biodiesel. Therefore, it is assumed that the oil diesel price is constant in all the simulations.

2.5.1 Counterfactuals

Given assumptions **S1** and **S2**, the diesel fuel price paid by the consumers (BX price) in a given month can be written as:

$$p_{it} = \alpha b_{it} + (1 - \alpha) d_{it} + \mu^r_{it} + \mu^w_{it} \tag{2.5.1}$$

Where b_t stands for the biodiesel producers' price in period t, d_t stands for the oil diesel producers' price in state i in period t, $\mu^r_{,i,t}$ is the retailers' markup in state i in period t, $\mu^w_{,i,t}$ is the wholesalers' markup in state i in period t, and α is the biodiesel mandate.

Equation 2.5.1 assumes that the one price law holds for biodiesel in all Brazilian territory and at regional level for oil diesel. Also, it implies that the fuel is not commercialized through auctions but directly negotiated between buyers and sellers.

The elements necessary to perform the simulation are given by equations 2.3.2, 2.4.5 and 2.5.1. Basically the exercise consists in finding a market equilibrium price for each biodiesel mandate keeping the other factors constant. For a given level of markup, oil diesel price and α, equation 2.5.1 gives the fuel diesel price as a function of the biodiesel price. Therefore, for the market equilibrium it is necessary to reach a biodiesel price that equates the quantity supplied (equation 2.4.5) and the quantity demanded (equation 2.3.2).

The computational details are as follows. First, I used just a month in the simulation to minimize possible dynamic distortions. The variables fleet, ICMS and petroleum diesel prices are those observed in November 2009. Second, the markup is the average margin for the period from January 2003 to December 2007 for the retailers and between January 2007 and December 2007 for the wholesalers. I used only one year to estimate the wholesalers' margin to capture its increase observed in the data. Finally, the coefficients for the demand are those estimated through fixed effects with IV (Table 2.4).

The algorithm used for the simulations is simple. It starts with a low biodiesel price and calculates the demand and supply for this price. If the difference between supply and demand is high, the price increases by a small amount and new values are calculated for demand and supply. The algorithm continues until the difference between the quantity supplied and the quantity demanded becomes negligible.

2.5.2 Results

The simulation results are shown in Table 2.6. The chart presents the scenarios with different mandatory percentages of biodiesel. The baseline is the scenario without biodiesel (100% of petroleum diesel). The second scenario has a 5% biodiesel mandate. For the other scenarios, the mandate was increased by 1% each scenario, up to a mandate of 15%, and a final simulation was run for a mandate of 20%.. In all simulations the possible loss of efficiency in the diesel engines, due to the increase of the proportion of biodiesel, is not considered.

It should be noted that the biodiesel price presented in the table corresponds to the producer price, free of wholesalers' and retailers' margins. The BX price, on the other hand, is the consumer's final price considering both margins and the proportion of biodiesel and petroleum diesel in the fuel. Furthermore, the BX price is the Brazilian average price. Since equation 2.3.2 allows for state effects, an equilibrium price and quantity is obtained for each state and the equilibrium quantity is used as a weight to calculate the national average price.

Table 2.6: Simulation Results

	Biodiesel		BX		Capacity
%	Price *	Quantity**	Price*	Quantity**	Utilization (%)
0%	-	-	1.75	3,459.70	0%
5%	1.57	170.5	1.78	3,409.20	44%
6%	1.58	203.9	1.78	3,398.70	52%
7%	1.58	237.2	1.79	3,388.40	61%
8%	1.59	270.2	1.80	3,377.90	70%
9%	1.59	303.1	1.80	3,367.60	78%
10%	1.59	335.7	1.81	3,357.10	86%
11%	1.59	368.2	1.97	3,347.30	95%
12%	2.14	388.5	2.07	3,237.20	100%
13%	3.47	388.5	2.09	2,988.20	100%
14%	4.62	388.5	2.26	2,774.70	100%
15%	5.64	388.5	2.45	2,589.70	100%
20%	9.33	388.5	3.58	1,942.30	100%

Notes: (*)R$ Jan 2003/L, (**) 1000 m^3. Values obtained through simulation based on equilibrium conditions. The first line (B0) presents the petroleum diesel price.

One can see that B5 increases the consumer price by 1.7%, from R$ 1.75 to R$ 1.78. The raise in price leads to a reduction in consumption of about 50.000 m^3, around 1.5%. If one doubles the current biodiesel mandate, from 5% to 10%, the price effect is more than proportional: fuel diesel price increases 3.6% reaching R$ 1.82. Consumption, on the other hand, decreases by 3% due to the use of B10.

The simulations also allow understanding of relations among capacity utilization, price and consumption. The increase in the capacity utilization from 44% (with B5) to 95% (with B11) increases the producer price by R$ 0.02 (about 1%). Therefore, the supply curve of the industry seems to be smooth on price up to 95% of its full capacity.

As stated before, B5 drives the capacity utilization to 44%. When one doubles the compulsory percentage, the capacity utilization goes to 86%. This less than proportional increase in the capacity utilization is due to the increase in price that reduces the consumption of the total mixture.

Further increases in the compulsory proportion of biodiesel increase the capacity utilization and lead to operation of firms with higher marginal costs. The effects on prices are higher after 11% of biodiesel. The biodiesel price increases 34% when B11 is substituted for B12.

It happens because with a 12% biodiesel mandate the industry starts to work to its full capacity, the price need to go up to equalize the demand and the offer. Since biodiesel is a small proportion of the total BX, its price needs a great increase to generate a significant drop in the BX demand. At a mandatory 20% blend, the price to the consumer is more than the double of the pure petroleum diesel, and the total consumption falls about 44%.

The impacts of different diesel mandates on consumer surplus, and in the profits[14] are also analyzed. The focus is the difference between the baseline scenario (no biodiesel) and the alternative scenarios with different percentages of biodiesel. Also, as there is no closed form for the consumers' utility, it is not possible to analyze the equivalent and the compensating variations. Alternatively, consumer surplus[15] is used as a measure of consumers' loss of welfare due to the change in price of diesel fuel caused by the biodiesel mandates. Finally, oil refineries' profits were not considered as no information regarding their costs is available. The results are summarized in table 2.7.

The second column in table 2.7 shows the total consumer surplus (the sum over Brazilian states). The current proportion of biodiesel (B5) causes a loss of R$ 104 million to the consumers. An increase to 10% in the mandatory percentage of biodiesel would cost R$ 212 million to consumers in terms of welfare. The loss increases very fast after 11% and achieves the impressive number of R$ 2 billion with B15 and R$ 4 billion with B20.

For a proportion of biodiesel up to 10%, the biodiesel industry's profits are quite small when compared to the consumer loss. At the current percentage (5%) the profits of all producers are around R$ 5 million. The double of the present percentage of biodiesel would increase the profits to 10 million. After that, however, one can observe a remarkable increase in the total profits; they achieve R$ 3.3 billion when one considers the scenario with 20% of biodiesel. The retailers and wholesalers are also affected by the introduction of biodiesel. Together they have a reduction of R$ 39 million in total profit at the current level of biodiesel. At a 20% level of biodiesel the total profit loss to retailers and wholesalers is above one billion.

Altogether, the losses to consumers, wholesalers and retailers outweigh the profits for the biodiesel producers in any level of biodiesel proportion considered. However, other benefits created by the use of biodiesel could make the adoption of biodiesel a socially optimal

[14]For biodiesel producers, Fixed costs are not considered. Therefore, profit is equal to revenue (price times quantity) minus marginal costs times quantity. If one consider the presence of fixed costs, the profit in the table becomes producer surplus. For retailers and wholesalers the variations in profits are equal to the margin times the change in the quantity sold.

[15]See Varian (1992).

Table 2.7: Simulation Results

Biodiesel Mandate	Consumer Surplus(*)	Biodiesel Producers' Profit(*)	Δ Retailers' Profit(*)	Δ Wholesalers' Profit(*)
5%	-104	5	-9	-30
6%	-125	7	-11	-36
7%	-146	8	-13	-42
8%	-168	9	-15	-49
9%	-190	10	-16	-55
10%	-212	10	-18	-61
11%	-233	254	-20	-67
12%	-470	801	-40	-133
13%	-1042	1272	-84	-281
14%	-1579	1682	-122	-408
15%	-2077	2044	-155	-519
20%	-4241	3365	-271	-905

Notes: (*)R$ Jan 2003/L. Results obtained by simulation.

decision. For example, environmental benefits, such as reduction of greenhouse gas emissions, or improvement of air quality may generate a significant welfare. Also, the income transfer to small farmers might be socially desirable. These considerations are above the scope of this paper and may be included in future work.

2.6 Conclusions, Limitations, and Future Research

The compulsory adoption of biodiesel in the Brazilian fuel market can bring many changes to the economy. Most of these changes are, however, still unknown. This paper addresses market equilibrium outcomes: the effects on price, consumption and welfare in the short run.

The analysis of the demand shows that even though diesel is an important raw material in the transport and agriculture industries (with almost no substitute in the short run) it has a considerable high price elasticity when controlled by other factors.

The industry supply estimation was based on the results of the last two auctions in the dataset. This simplification was adopted to avoid capturing possible dynamic effects on the static analysis done in the paper. The supply curve constructed is considerably smooth up to the industry total capacity. After that, a change in the compulsory proportion has a huge impact on prices due to capacity constraints.

The simulations show that the current level of biodiesel raises the final price to consumers by 1.7% and decreases the consumption by 1.5%. If the government doubles the compulsory proportion the new price would be 3.5% higher with a decrease in consumption of 3%. Additionally, the welfare analysis showed that the consumers, wholesalers and retailers have a huge loss in consumer surplus and in profits due to the adoption of the biodiesel. With a 5% biodiesel blend, the total loss is around R$ 143 million.

The results obtained in this study should be taken with caution as they depend on the set of assumptions. Some assumptions may be considered controversial and they are required for two reasons. First, the complexity involved in the adoption of biodiesel (and biofuels in general). It is impossible for a quantitative study to deal with all the aspects involved in a biofuel program, which include a large supply chain and political questions. Second, some hypotheses are necessary due lack of data. For example, no data on the individual decisions taken by farmers is available. The paper, therefore, should be seen as an indication of the costs involved in the biodiesel program. More studies are necessary for a better understanding of the overall impact.

The political implications of the program are not considered in this paper. The price increase observed in the simulations could not be politically acceptable. Therefore, the price formation assumption (free negotiation) might not hold. Also it is unknown how the increase in the oil production and oil refining in Brazil would affect the biofuel programs (the country is a net exporter of oil and may become a net exporter of petroleum diesel). This increase could change the government support to the biodiesel program.

The paper also does not consider the program effects in all the parts of the supply chain. This includes the effects of biodiesel in feedstock, land use, employment and the expenditure in subsidies. For, example the alternative uses of the raw materials involved in the biodiesel production must affect the biodiesel price, and consequently the final blended diesel price. Future work, should look with attention to this important part of the supply chain.

Furthermore, the paper did not apply a dynamic framework. Future work should improve the analysis in this aspect in order to create more realistic scenarios. Combined, all this information could be a better guide for energy policy decisions.

Bibliography

[1] Almeida, E. F. d., Bomtempo, J. V., Silva, C. M. d. S. e., (2008). “Biofuels: Linking Support to Performance.” OECD.

[2] Althoff, K., Ehmke, C., Gray, A. W., 2003.“Economic analysis of alternative indiana state legislation on biodiesel. Tech. rep., Center for Food and Agricultural Business, ” Department of Agricultural Economics

[3] Althoff, K., Ehmke, C., Gray, A. W., (2003). “Economic analysis of alternative indiana state legislation on biodiesel.” *Technical Report, Center for Food and Agricultural Business, Department of Agricultural Economics, Pardue University.*

[4] ANP, (2010). National agency for oil and biofuels. www.anp.gov.br, [In Portuguese].

[5] demirbas, D., (2007). “Importance of biodiesel as transportation fuel.” *Energy Policy*, **35(9)**, 4661-4670.

[6] Barbosa, M. Z.,(2011). “Disponibilidade de oleo de soja para energia e alimentos no Brasil.” *Informações Econômicas*, **41(9)**.

[7] Basso, L. J., Oum, T. H.,(2007).“Automobile fuel demand: A critical assessment of empirical methodologies.” *Transport Reviews*, **27(4)**, 449-484.

[8] Brazil, November, 14 (1975). “Law number 76.593”Ḋiério Oficial [da] República Federativa do Brasil.

[9] Brazil, November, 14 (1976). “Law number 76.593”Ḋiário Oficial [da] República Federativa do Brasil.

[10] Brazil, August, 6 (1997). “Law number 9.478”Ḋiário Oficial [da] Rep
’ublica Federativa do Brasil.

[11] Brazil, January, 13 (2005). "Law number 11.097"Ḋiário Oficial [da] República Federativa do Brasil.

[12] Dahl, C., Sterner, T., (1991). "Analysing gasoline demand elasticities." *Energy Economics*, **3(13)**, 203-210.

[13] De Gorter, H., Just, D. R., (2009). "The economics of a blend mandate for biofuels." *American Journal of Agricultural Economics*, **91 (3)**, 738-750.

[14] Dixit, A., Stiglitz, J. E., (1977). "Monopolistic competition and optimum product diversity."
textitAmerican Economic Review.

[15] Donald, S., Paarsch, H., (1993)."Identifcation, estimation, and testing in empirical models of auction within the independent private values paradigm. " *UWO Department of EconomicsWorking Papers 9319, University of Western Ontario, Department of Economics.*

[16] Ericson, R., Pakes, A., (1995)."Markov-perfect industry dynamics: A framework for empirical work." *The Review of Economic Studies*, **62(1)**, 53-82.

[17] Fargione, J., Hill, J., Tilman, D., Polasky, S., Hawthorne, P., (2008). "Land clearing and the biofuel carbon debt." *Science*, **319 (5867)**, 1235-1238.

[18] Greene, D., (1982). "State level stock system of gasoline demand." *Transportation Research Record*, **(801)**, 44-51.

[19] Greene, W. H., (2003). *Econometric Analysis*. Prentice Hall.

[20] Guerre, E., Perrigne, I., Vuong, Q., (2000). "Optimal nonparametric estimation of first-price auctions." *Econometrica*, **68(3)**, 525-574.

[21] Hosken, D. S., McMillan, R. S., Taylor, C. T., (2008). "Retail gasoline pricing: What do we know?" *International Journal of Industrial Organization*, **26(6)**, 1425-1436.

[22] Jia, P., 2006. "Estimating the effects of global patent protection in pharmaceuticals: A case study of quinolones in india." *American Economic Review*, **96(5)**, 1477-1514.

[23] Krugman, P., (1979). "Increasing returns, monopolistic competition and international trade". *Journal of International Economics*, **9** 469-479.

[24] Laffont, J.-J., Ossard, H., Voung, Q., (1995). "Econometrics of first-price auctions". *Econometrica*, **63(4)**, 953-980.

[25] Lamers, P., McCormick, K., Hilbert, J. A., (2008). "The emerging liquid biofuel market in Argentina: Implications for domestic demand and international trade". *Energy Policy*, **36(4)**, 1479-1490.

[26] Li, T., Zheng, X., (2009)."Entry and competition effects in first-price auctions:theory and evidence from procurement auctions." *Review of Economic Studies*, **76(4)**, 1397-1429.

[27] Mas-Colell, A., Whinston, M. D., Green, J. R., (1995). *Microeconomic Theory.* Oxford University Press.

[28] Maskin, E., Tirole, J., (1988)."A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs." *Econometrica*, **56(3)**, 549-569.

[29] MME, (2005). Ministerio das minas e energia. "Programa Nacional de Produção e Uso do Biodiesel," available in: http://www.mme.gov.br/programas/biodiesel/. [In Portuguese].

[30] Nevo, A., (2001). "Measuring Market power in the ready-to-eat cereal industry." *Econometrica*, **69(2)**, 307-342.

[31] Pfuderer, S., Davies, G., Mitchell, I., (2009)."The 2007/08 agricultural price spikes: Causes and policy implications." *Technical Report*, HM Government.

[32] Ryan, S. P. (2012): "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, **80**, 1019-1061.

bibitemSilva César, A. d., Batalha, M. O., (2010)."Biodiesel in brazil: History and relevant policies." *African Journal of Agricultural Research*, **5(11)**.

[33] Stock, J. H., Wright, J. H., Yogo, M., October (2002). "A survey of weak instruments and weak identification in generalized method of moments." *Journal of Business & Economic Statistics*, **4(20)**, 518-529.

[34] Sutton, J., (1991). *Sunk Costs and Market Structure*. Cambridge, MA: MIT Press.

[35] Tirole, J., (1988). *The Theory of Industrial Organization*. MIT University Press.

[36] Varian, H. R., (1992). *Microeconomic Analysis*, Third Edition. W. W. Norton & Company.

[37] Wooldridge, J. M., (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT University Press.

Appendix

Table 2.8: Groups 1, 2, and 3

Plant	Monthly capacity(*)	Marginal Cost(**)	Group
Biolix	900	1592.92	1
Brasil ecodiesel 1	9000	1592.92	1
Brasil ecodiesel 2	8100	1592.92	1
Coomisa	360	1592.92	1
Cooperfeliz	200	1592.92	1
Ouro verde	510	1592.92	1
Abdiesel	72	1592.92	2
Agrenco	19608	1592.92	2
Big Frango	1200	1592.92	2
Bionorte	2451	1592.92	2
Cooperbio	120	1592.92	2
Fusermann	900	1592.92	2
Granol	7500	1592.92	2
Grupal	300	1592.92	2
Kgb	150	1592.92	2
Nutec	72	1592.92	2
Rondobio	300	1592.92	2
Soyminas	1200	1592.92	2
Taua	3000	1592.92	2
Tecnodiesel	330	1592.92	2
Usibio	600	1592.92	2
Vermoehlen	150	1592.92	2
Araguassu	3000	1533.34	3
Bio Oleo	300	1494.74	3
Brasil Ecodiesel 3	10800	1543.09	3
Clv	3000	1570.53	3
Innovatti	900	1552.78	3
Ssil	150	1396.85	3

Notes: (*) m^3,(**)R$ Jan 2003/L. VElaborated based on data from the National Agency for Oil and Biofuels ANP (2010).

Table 2.9: Group 4

Plant	Monthly capacity(*)	Marginal Cost(**)	Group
Abdiesel	180	1402.74	4
Adm	28650	1543.09	4
Agropalma	900	1588.78	4
Agrosoja	2400	1581.55	4
Amazonbio	1350	1402.74	4
B-100	900	1354.39	4
Barralcool	4902	1577.34	4
Beira Rio	360	1400.71	4
Binatural	9000	1480.71	4
Biocamp	4620	1588.78	4
Biocapital	24720	1581.55	4
Biocar	900	1470.53	4
Biopar parana	3600	1583.86	4
Biopar parecis	700	1590.11	4
Biotins	810	1332.92	4
Bioverde	7353	1579.65	4
Bracol	16807	1568.43	4
Brasil Ecodiesel 4	10800	1591.51	4
Brasil ecodiesel 5	13320	1587.30	4
Brasil ecodiesel 6	10800	1592.92	4
Bsbios	13320	1578.95	4
Caramuru	18750	1573.69	4
Cesbra	1800	1396.85	4
Comanche	10050	1535.51	4
Cooperbio 1	10200	1586.18	4
Cooperbio 2	120	1403.86	4
Dvh	1050	1403.86	4
Fertibom	4200	1563.02	4
Fiagril	12299	1578.95	4
Granol 1	18390	1590.11	4
Granol 2	28000	1573.69	4
Oleoplan	19800	1587.30	4
Petrobras 1	9051	1568.43	4
Petrobras 2	9051	1568.43	4
Petrobras 3	9051	1543.79	4
Sp Bio	2082	1400.14	4
Transportadora Caibiense	3000	1402.74	4

Notes: (*) m^3,(**)R$ Jan 2003/L. Elaborated based on data from the National Agency for Oil and Biofuels ANP (2010).

Chapter 3

An Alternative Asymptotic Least Squares Estimator for Dynamic Games

Co-authored with Sorawoot Srisuma (University of Surrey) and Fabio A. Miessi Sanches (University of São Paulo)

Abstract

The estimation of dynamic games is known to be a numerically challenging task. In this paper we propose an alternative class of asymptotic least squares estimators to Pesendorfer and Schmidt-Dengler's (2008), which includes several well known estimators in the literature as special cases. Our estimator can be substantially easier to compute. In the leading case with linear payoffs specification our estimator has a familiar OLS/GLS closed-form that does not require any optimization. When payoffs have partially linear form, we propose a sequential estimator where the parameters in the nonlinear term can be estimated independently of the linear components, the latter can then be obtained in closed-form. We show the class of estimators we propose and Pesendorfer and Schmidt-Dengler's are in fact asymptotically equivalent. Hence there is no theoretical cost in reducing the computational burden. Our estimator seem to perform well in a simple Monte Carlo experiment.

3.1 Introduction

We consider the estimation problem for a class of dynamic games of incomplete information that generalizes the single agent discrete Markov decision models surveyed in Rust (1994); for a recent survey see Aguirregabiria and Mira (2010). The setup is in an infinite time horizon, where players' private values enter the payoff function additively and are independent across players, under the conditional independence framework. A Markov equilibrium of such game can be represented by a fixed point of nonlinear equations in the space of choice probabilities and has been shown to exist (e.g. see Aguirregabiria and Mira (2007) and Pesendorfer and Schmidt-Dengler (2008)). A variety of methods have been proposed by different authors to estimate the same class of games based on the equilibrium condition in recent years; examples are given below. However, a common component of these methodologies is a nonlinear optimization problem that may act as a considerable deterrent for applied researchers to estimate dynamic games due to involved programming needs and/or long computational time.

In this paper we propose a class of asymptotic least squares estimators constructed based on the equilibrium condition of the game when represented in the space of payoffs. Our work is motivated by the well-received methodology developed in Pesendorfer and Schmidt-Dengler (2008), who propose an efficient estimator and provide a unifying framework that includes the non-iterative pseudo-likelihood estimator of Aguirregabiria and Mira (2007) and the moment based estimators discussed in Pakes, Ostrovsky and Berry (2007) as special cases. In contrast to ours, Pesendorfer and Schmidt-Dengler use the choice probability representation of the equilibrium to construct their estimator. Our goal is to show there is much to gain computationally using our approach at no cost. Henceforth we use the abbreviation ALSE_{PSD} when referring to a generic estimator of Pesendorfer and Schmidt-Dengler.

We claim our estimator can be substantially easier to compute than ALSE_{PSD}. In the leading case our estimator has a familiar OLS/GLS closed-form expression when the per-period payoff function takes a linear-in-parameter specification.[1] In an intermediate case

[1] The linear payoffs structure may seem restrictive, but it is in fact quite general as it includes any nonlinear (basis) functions of observables; albeit perhaps with an atheoretic flavor. However, linear specification arises naturally in many applications, and/or does not cause much concern in terms of structural interpretability in other situations. A leading example for the latter is when the goal of an empirical analysis is to study market outcomes, such as competition study of market power. Some notable recent empiri-

when the payoff function has an additive partially linear form, Frisch-Waugh-Lovell theorem can be applied so the parameters in the nonlinear part can be estimated first (dimensional reduction), and the linear-in-parameter component can be obtained in closed-form in the second step.[2] Even in a more general nonlinear case, we argue that our estimator is still generally easier to compute than ALSE_{PSD}. ALSE_{PSD} also provides a good benchmark for a comparison with other estimators in the literature as it has a well-defined efficiency property. We establish a duality between our estimator and ALSE_{PSD}, in the sense that they can always be constructed to have the same asymptotic distribution. Therefore our efficient estimator is as efficient as the efficient ALSE_{PSD}.

The large sample properties of our estimator (and for asymptotic least squares generally) are easy to derive for discrete games. Technically, our estimation problem is a least squares problem with generated regressors and regressands, which are generally smooth functions of the finite dimensional first stage parameters that are nonparametrically identified. In addition, the number of square terms in the objective function does not grow with sample size but is determined by the cardinality of the action and state spaces. Therefore our estimator belongs to the class of asymptotic least squares estimators as defined in Gourieroux and Monfort (1985,1995) in the same sense as ALSE_{PSD}. The close connection between our estimator and ALSE_{PSD} goes even further given the smooth bijective relation between normalized expected payoffs and choice probabilities (Hotz and Miller (1993)'s inversion); ALSE_{PSD} is defined to minimize the distance between the probabilities implied by the pseudo-model and the data. We show that, locally around the true, using the inverse function theorem, our estimator can be constructed to have the same asymptotic distribution as any ALSE_{PSD} by choosing an appropriate weighting matrix and vice versa.

There are at least two reasons why the estimation of dynamic games can be non-trivial. First, as well-known from the single-agent problem, it involves value functions that generally do not have closed-form and need to be numerically evaluated so it is computationally demanding (see Rust (1996)). For games, there is also a potential issue of indeterminacy of multiple equilibria that gives rise to incomplete models (Tamer (2003)). A novel approach popularized by Hotz and Miller (1993) performs inference on the pseudo-model, generated

cal applications of linear-in-parameter payoffs include Aguirregabiria and Mira (2007), Ryan (2012) and Collard-Wexler (2013).

[2] Modeling of additive linear components in the payoffs often appear in games with entry/exit decisions, as fixed cost or scrap value, or more generally as fixed effects.

from to the observed data, by estimating the (policy) value functions that can significantly simplify the computational aspect. Pseudo-models are also generally easier to handle in a strategic environment as they have been shown to be complete for several classes of games (Srisuma (2013)). Methodologies based on pseudo-models are often referred to as two-step estimators since they require estimation of value functions in the first stage. Many recently proposed estimators for dynamic games are two-step estimators.

However, despite the simplification of two-step methods, the numerical aspects for implementing existing estimators in the literature appear to remain a concern as they generally involve solving highly nonlinear optimization problems. It is not uncommon to see methodology papers using estimation time, amongst other things, as a competing factor. Furthermore, it is also not unusual that the choice of players' per-period payoff specification is chosen with the ease of numerical implementation in mind. In particular there can be substantial benefits (in terms of computational time) in specifying player's payoff functions to be linear-in-parameters. As the action-specific expected payoffs can then be written as a linear transformation of the parameter, following from the linear structure that defines the expected payoffs using stationary Markovian beliefs; examples of such discussions can be found in Bajari, Benkard and Levin (2007, Section 3.3.1) and Pakes, Ostrovsky, Berry (2007, Section 3). As a result, a linear parameterization of the payoffs is a leading specification employed in empirical work (see Footnote 1 for examples).

The objective functions that are used to define many two-step estimators in the literature are constructed in terms of choice probabilities implied by the pseudo-model. These probabilities can be motivated by the equilibrium condition of the game, which can be stated in terms of consistent beliefs with probabilities of best responses. Choice probabilities are used to define traditional criterion functions such as pseudo-likelihood function (Aguirregabiria and Mira (2007), Kasahara and Shimotsu (2012)) or moment and minimum distance based conditions (Pakes, Ostrovsky, Berry (2007), Pesendorfer and Schmidt-Dengler (2008)). However, in order to calculate the probabilities implied by the pseudo-model, one must first compute the expected discounted payoffs that determine the region of integration to be integrated to compute the probabilities. Furthermore, the integral is generally a nonlinear map of the expected payoffs, and it typically has to be computed numerically outside the well-known conditional logit framework. The integral, following Hotz and Miller (1993)'s inversion result, in fact represents a one-to-one mapping between the probabilities and the normalized expected payoffs.

There are also other methodologies that use expected payoffs explicitly to define their objective functions. The first such two-step estimator has been developed by Hotz, Miller, Sanders and Smith (1994), who estimate the expected payoffs by forward simulation, to estimate a dynamic decision problem for a single agent. Hotz et al. define their estimator using conditional moment restrictions. They also recognize it is possible to have a closed-form estimator when payoff functions have linear-in-parameter specification in the form of an IV estimator (see equation (5.8) in the Monte Carlo Study section of Hotz et al. (1994)). In the context of dynamic games we are only aware of two other current methodologies that base their objective functions explicitly on expected payoffs. First is the two-step estimator proposed by Bajari, Benkard and Levin (2007), who also use forward simulation like Hotz et al. However, generally no closed-form estimator is possible with Bajari, Benkard and Levin's methodology as they compare expected payoffs in the pseudo-model and those generated by local perturbations. The other is Bajari, Chernozhukov, Hong and Nekipelov (2009), who provide nonparametric identification results for a more general game with continuous state space and propose an efficient one-step estimator.[3,4]

The rest of the paper is organized as follows. Section 3.2 begins with an illustrative example that motivates our estimator, and then describes the model and our estimator for games. Section 3.3 gives the main results. Section 3.4 presents results from Monte Carlo experiments that compare the statistical performance and relative speed of our estimator compared to ALSE_{PSD}. Section 3.5 apted or applied to complement other recent results in the literature. All proofs can be found in the Appendix.

3.2 Methodology

We begin with an illustration that highlights the idea behind computational advantages of our estimation approach. Section 3.2.1 describes elements of the game. We define the pseudo-model in Section 3.2.2 and introduce our estimator in Section 3.2.3.

[3]An earlier version of Bajari et al. (2009), Bajari and Hong (2006), proposes a two-step estimator that can be seen as the dynamic game version of Hotz et al. (1994).

[4]Another notable estimator that does not take a two-step approach is Egesdal, Lai and Su (2012). However, Egesdal et al. construct their objective functions in terms of choice probabilities.

3.2.1 Least Squares in Probabilities vs Payoffs

Consider a model generated by the following binary choice variable:

$$a_t(\theta) = \mathbf{1}\left[v_\theta(x_t) \leq \varepsilon_t\right] \quad \text{for} \quad \theta \in \Theta \subset \mathbb{R}^p,$$

where x_t and ε_t are independent. Let the cdf of ε_t be denoted by Q. For all x, let $P_\theta(x) = \Pr\left[a_t(\theta) = 1 | x_t = x\right]$, so that $P_\theta(x) = Q(v_\theta(x))$. Assume the support of x_t is finite, say $\{x^j\}_{j=1}^J$ for some $J < \infty$, so that we can define $\mathbf{P}_\theta = \Gamma(\mathbf{v}_\theta)$, where $\mathbf{P}_\theta = (P_\theta(x^1), \ldots, P_\theta(x^J))^\top$, $\mathbf{v}_\theta = (v_\theta(x^1), \ldots, v_\theta(x^J))^\top$ and $\Gamma(\mathbf{v}_\theta) = (Q(v_\theta(x^1)), \ldots, Q(v_\theta(x^J)))^\top$.

Suppose: we observe a random sample of $\{a_t, x_t\}$ where $a_t = a_t(\theta_0)$ for some $\theta_0 \in \Theta$, which is the parameter value of interest; v_θ is nonparametrically identified up to θ, and there exists a consistent estimator of $\mathbf{v}_\theta$, say $\widehat{\mathbf{v}}_\theta$, for all θ; and, Q is known and invertible. Let $\mathbf{P} = (P(x^1), \ldots, P(x^J))^\top$ be a vector of choice probabilities identified from the data, so that $\mathbf{P} = \mathbf{P}_{\theta_0}$, then one may consider a class of estimators defined by

$$\widehat{\theta}_p(\mathcal{V}) = \arg\min_{\theta \in \Theta} \left(\widetilde{\mathbf{P}} - \widehat{\mathbf{P}}_\theta\right)^\top \mathcal{V} \left(\widetilde{\mathbf{P}} - \widehat{\mathbf{P}}_\theta\right), \tag{3.2.1}$$

where $\widetilde{\mathbf{P}}$ and $\widehat{\mathbf{P}}_\theta$ are estimators for $\mathbf{P}$ and $\mathbf{P}_\theta$ respectively, and $\mathcal{V}$ be some positive definite matrix. Note that $\widetilde{\mathbf{P}}$ and $\widehat{\mathbf{P}}_{\theta_0}$ are generally different since the former is model-free while the latter is estimated through $\widehat{\mathbf{v}}_\theta$. Similarly, we can define $\mathbf{v} = (Q^{-1}(P(x^1)), \ldots, Q^{-1}(P(x^J)))^\top$, which is also identified from the data, so that $\mathbf{v} = \mathbf{v}_{\theta_0}$ by construction. Then one can also consider an alternative class of estimators:

$$\widehat{\theta}_v(\mathcal{W}) = \arg\min_{\theta \in \Theta} \left(\widetilde{\mathbf{v}} - \widehat{\mathbf{v}}_\theta\right)^\top \mathcal{W} \left(\widetilde{\mathbf{v}} - \widehat{\mathbf{v}}_\theta\right), \tag{3.2.2}$$

where $\widetilde{\mathbf{v}}$ is $\Gamma^{-1}(\widetilde{\mathbf{P}})$ and $\mathcal{W}$ is a positive definite matrix. As described previously, $\widetilde{\mathbf{v}}$ and $\widehat{\mathbf{v}}_{\theta_0}$ will also generally differ.

Equations (3.2.1) and (3.2.2) provide two different estimators for θ_0. We argue the latter should generally be easier to compute than the former since it is more convenient to compute $(\widetilde{\mathbf{v}}, \widehat{\mathbf{v}}_\theta)$ relative $(\widetilde{\mathbf{P}}, \widehat{\mathbf{P}}_\theta)$ across different values of θ. This argument is most transparent when v_θ has a linear-in-parameter specification, i.e. $v_\theta(x_t) = \theta^\top v(x_t)$ for some $p-$dimensional vector $v(x_t)$. Then $\widehat{\mathbf{v}}_\theta$ can be written as $\widehat{\mathbf{X}}\theta$, where $\widehat{\mathbf{X}}$ is a J by p matrix such that its $j-$th row equals $\widehat{v}(x^j)^\top$. The solution to (3.2.2) is unique and has a closed-form, $\left(\widehat{\mathbf{X}}^\top \mathcal{W} \widehat{\mathbf{X}}\right)^{-1} \widehat{\mathbf{X}}^\top \mathcal{W} \widehat{\mathbf{v}}$, when $\widehat{\mathbf{X}}^\top \mathcal{W} \widehat{\mathbf{X}}$ is invertible. Even without the linear parameterization of v_θ, every evaluation of $\widehat{\mathbf{P}}_\theta$ requires the mapping of $v_\theta(x^j)$ by Q for all j, for every θ, where Q is generally a nonlinear function that may have to be computed numerically. In contrast, for (3.2.2), the potentially costly step of applying Q^{-1} has to be performed only once to estimate $\mathbf{v}$ that does not depend on θ. Regardless of the parameterization in v_θ, under some suitable regularity conditions, and appropriate choices of weighting matrices, the two estimators can be shown to be asymptotically equivalent near θ_0 in the sense that there exists $\mathcal{W}_\mathcal{V}$ and $\mathcal{V}_\mathcal{W}$ such that for any $\mathcal{V}$ and $\mathcal{W}$:

$$
\begin{aligned}
\sqrt{N}\left(\widehat{\theta}_v(\mathcal{W}_\mathcal{V}) - \theta_0\right) &= \sqrt{N}\left(\widehat{\theta}_p(\mathcal{V}) - \theta_0\right) + o_p(1), \\
\sqrt{N}\left(\widehat{\theta}_p(\mathcal{V}_\mathcal{W}) - \theta_0\right) &= \sqrt{N}\left(\widehat{\theta}_v(\mathcal{W}) - \theta_0\right) + o_p(1),
\end{aligned}
$$

where N denotes the sample size.

The estimator in (3.2.1) is closely related to ALSE_{PSD} and other Hotz and Miller (1993)'s type estimators that have been widely adopted in the dynamic game setting. In contrast the estimator based on (3.2.2) is the asymptotic least squares analog to the estimator proposed in Hotz et al. (1994). For the remainder of this section we develop an estimator based on (3.2.2) in the context of a dynamic game.

3.2.2 Framework

We consider a game with I players, indexed by $i \in \mathcal{I} = \{1, \ldots, I\}$, over an infinite time horizon. The elements of the game in each period are as follows:

Actions. For notational simplicity we assume all players have the same action space. The action set of each player is $A = \{0, 1, \ldots, K+1\}$. We denote the action variable for player i by a_{it}. Let $\mathbf{a}_t = (a_{1t}, \ldots, a_{It}) \in \mathbf{A} = \times_{i=1}^I A$. We will also occasionally abuse the notation and write $\mathbf{a}_t = (a_{it}, \mathbf{a}_{-it})$ where $\mathbf{a}_{-it} = (a_{1t}, \ldots, a_{i-1t}, a_{i+1t} \ldots, a_{It}) \in \mathbf{A} \backslash A$.

States. Player i's information set is represented by the state variables $s_{it} \in S$, where $s_{it} = (x_{it}, \varepsilon_{it})$ such that $x_{it} \in X$ is common knowledge to all players and $\varepsilon_{it} \in \mathcal{E} = \mathbb{R}^{K+1}$ denotes private information only observed by player i. Note that common state space X is without any loss of generality. We shall use s_{it} and (x_t, ε_{it}) interchangeably. We define $(\mathbf{s}_t, \mathbf{s}_{-it}, \varepsilon_t, \varepsilon_{-it}, \mathcal{E})$ analogously to $(\mathbf{a}_t, \mathbf{a}_{-it}, A)$, and denote the support of $\mathbf{s}_t$ by $S = X \times \mathcal{E}$.

State Transition. Future states are uncertain. Players' actions and states today affect future states. The evolution of the states is summarize by a Markov transition law $P\left(\mathbf{s}_{t+1}|\mathbf{s}_t, \mathbf{a}_t\right)$.

Per Period Payoff Functions. Each player has a payoff function, $u_i : A \times S \to \mathbb{R}$, which is time separable. The payoff function for player i can depend generally on $(\mathbf{a}_t, x_t, \varepsilon_{it})$ but not directly on ε_{-it}.

Discounting Factor. Future period's payoffs are discounted at the rate $\beta_i \in (0, 1)$ for each player. For notational simplicity we take $\beta_i = \beta$ for all i.

We impose the following assumptions throughout the paper.

Assumption M1 *(Additive Separability).* $u_{i,\theta_i}\left(a_i, \mathbf{a}_{-i}, x, \varepsilon_i\right) = \pi_{i,\theta_i}\left(a_i, \mathbf{a}_{-i}, x\right)$ $+ \sum_{a' \in A} \varepsilon_i\left(a'\right) \mathbf{1}\left[a_i = a'\right]$ *for all* $i, \theta_i, a_i, \mathbf{a}_{-i}, x, \varepsilon_i$, *where* π_{i,θ_i} *is known up to* $\theta_i \in \Theta_i \subset \mathbb{R}^{p_i}$.

Assumption M2 *(Conditional independence). The transitional distribution of the states has the following factorization:* $P\left(x_{t+1}, \varepsilon_{t+1}|x_t, \varepsilon_t, \mathbf{a}_t\right) = Q\left(\varepsilon_{t+1}\right) G\left(x_{t+1}|x_t, \mathbf{a}_t\right)$, *where* Q *is the cumulative distribution function of* ε_t *and* G *denotes the transition law of* x_{t+1} *conditioning on* $\mathbf{a}_t$ *and* x_t.

Assumption M3 *(Independent private values). The private information is independently distributed across players, and each is absolutely continuous with respect to the Lebesgue measure whose density is bounded on* $\mathbb{R}^{K+1}$. *So that* $Q\left(\varepsilon\right) = \prod_{i=1}^{I} Q_i\left(\varepsilon_i\right)$, *where* Q_i *denotes the cumulative distribution function of* ε_{it}.

Assumption M4 *(Discrete public values). The support of* x_t *is finite so that* $X = \left\{x^1, \ldots, x^J\right\}$ *for some* $J < \infty$.

M1 - M4 are standard in the modeling of dynamic discrete games in the literature. Note that M2 implies x_t and ε_t are independent, however, this can be relaxed slightly at the cost of more notation by changing all of our statements regarding Q and Q_i to be taken conditional on x_t. M4 is also not essential for the general idea behind estimation of dynamic games. Although the complexity of the asymptotic theory and the practical aspects increase significantly when x_t includes continuous random variables; see Bajari et al. (2009) and Srisuma and Linton (2012).

At time t every player observes s_{it}, each then chooses a_{it} simultaneously. We consider a Markovian framework where players' behaviors are stationary across time and players are assumed to play pure strategies. More specifically, for some $\alpha_i : S \to A$, $a_{it} = \alpha_i(s_{it})$ for all i, t, so that whenever $s_{it} = s_{i\tau}$ then $\alpha_i(s_{it}) = \alpha_i(s_{i\tau})$ for any τ. The beliefs are also time invariant. Player i's beliefs, σ_i, is a distribution of $\mathbf{a}_t = (\alpha_1(s_{1t}), \ldots, \alpha_I(s_{It}))$ conditional on x_t for some pure Markov strategy profile $(\alpha_1, \ldots, \alpha_I)$. The decision problem for each player is to solve

$$\begin{aligned} \max_{a_i \in A_i} \{ E_{\sigma_i}[u_{i,\theta_i}(a_{it}, \mathbf{a}_{-it}, s_i) | s_{it} &= s_i, a_{it} = a_i] + \beta E_{\sigma_i}[W_{i,\theta_i}(s_{it+1}; \sigma_i) | s_{it} = s_i, a_{it} = a_i] \} \\ \text{where } W_{i,\theta_i}(s_i; \sigma_i) &= \sum_{\tau=t}^{\infty} \beta^{\tau-t} E_{\sigma_i}[u_{i,\theta_i}(\mathbf{a}_\tau, s_{i\tau}) | s_{it} = s_i], \end{aligned} \tag{3.2.3}$$

for any s_i. The subscript σ_i on the expectation operator makes explicit that present and future actions are integrated out with respect to the beliefs σ_i; in particular, player i forms an expectation for all players' future actions including herself, and todays actions of opposing players. $W_{i,\theta_i}(\cdot; \sigma_i)$ is a policy value function since the expected discounted return needs not be an optimal value from an optimization problem since σ_i can be any beliefs, not necessarily equilibrium beliefs. Note that the transition laws for future states are completely determined by the primitives and the beliefs. Any strategy profile that solves the decision problems for all i and is consistent with the beliefs satisfies is an equilibrium strategy. It is well-known that players' best responses are pure strategies almost surely and Markov perfect equilibria for games under M1 - M4 (e.g. see Aguirregabiria and Mira (2007) and Pesendorfer and Schmidt-Dengler (2008)). However, there may be multiple equilibria.

3.2.3 Pseudo-Model

We now define the pseudo-model that plays a central role in two-step estimation methods. The starting point is the structural assumption that we observe random sample of $\{\alpha_1^*(s_{1t}), \ldots, \alpha_I^*(s_{It}), x_t, x_{t+1}\}$ from a single equilibrium, where $\alpha_i^* = \alpha_{i,\theta_{i0}}$ for some $\theta_{i0} \in \Theta_i \subset \mathbb{R}^{p_i}$ for all i. Let $P_i^*(a_i|x) = \Pr[\alpha_i^*(s_{it}) = a_i | x_t = x]$ for all a_i, x. Then we have: (i) the equilibrium beliefs for all players is summarized by $\prod_{i=1}^I P_i^*$; (ii) $\Pr[a_{it} = a_i | x_t = x] = P_i^*(a_i|x)$ and $\Pr[x_{t+1} = x' | x_t = x, a_t = a] = G(x'|x, a)$ for all a, x, x'. For notational simplicity, for this section and the next, we shall: omit *; let α_i and P_i denote the equilibrium strategy and choice probability function for player i; and, without any ambiguity let $a_{it} = \alpha_i(s_{it})$ for all i, t. Then the pseudo-model can be defined as a collection of joint conditional distributions indexed by $\theta = (\theta_1^\top, \ldots, \theta_I^\top)^\top \in \times_{i=1}^I \Theta_i = \Theta \subset \mathbb{R}^p$. Also let θ_0 denote $(\theta_{10}^\top, \ldots, \theta_{I0}^\top)^\top$.

Definition: The pseudo-model is $\{P_\theta\}_{\theta\in\Theta}$ such that $P_\theta = \prod_{i=1}^I P_{i,\theta_i}$ and for all i, θ_i, a_i, x:

$$\begin{aligned} P_{i,\theta_i}(a|x) &= \Pr\left[\alpha_{i,\theta_i}(s_{it}) = a | x_t = x\right] \quad \text{a.s.,} \quad \text{where} \\ \alpha_{i,\theta_i}(s_{it}) &= \arg\max_{a_i\in A}\left\{E\left[\pi_{i,\theta_i}(a_i, a_{-it}, x_t)\right| x_t\right] + \varepsilon_{it}(a_i) + \beta E\left[V_{i,\theta_i}(s_{t+1})\right| x_t, a_{it} = a_i]\right\}, \\ V_{i,\theta_i}(s_{it}) &= E[\pi_{i,\theta_i}(a_{it}, a_{-it}, x_t) + \sum_{a'=0}^K \varepsilon_{it}(a')\mathbf{1}\left[a_{it} = a'\right]|s_{it}] + \beta E\left[V_{i,\theta_i}(s_{it+1})\right| s_{it}\right]. \end{aligned}$$

By construction $P_{i,\theta_i} = P_i$ for all i when $\theta_i = \theta_{i0}$ for all i, and V_{i,θ_i} also equals $W_{i,\theta_i}(\cdot;\sigma_i)$ (as defined in (3.2.3)), when $\sigma_i = \prod_{j=1}^I P_j$. Let $v_{i,\theta_i}(a_i, x) = E\left[\pi_{i,\theta_i}(a_i, a_{-it}, x_t)\right| x_t = x] + \beta E\left[V_{i,\theta_i}(s_{t+1})\right| x_t = x, a_{it} = a_i]$, then we can write

$$P_{i,\theta_i}(a|x) = \Pr\left[v_{i,\theta_i}(a_i, x_t) + \varepsilon_{it}(a_i) > v_{i,\theta_i}(a_i', x_t) + \varepsilon_{it}(a_i') \text{ for all } a_i' \neq a_i | x_t = x\right], \quad (3.2.4)$$

which is familiar from the classical random utility model (e.g. see McFadden (1974)) with mean utility v_{i,θ_i}. The numerical advantage in working with the *pseudo*-model, as opposed to the *actual* model, is that v_{i,θ_i} is relatively straightforward to compute for different θ_i, since all expectations that define v_{i,θ_i} are calculated independent of θ_i; all with respect to $P(s_{t+1}|s_t, a_t)$ for all players that is equivalent to earlier notation using E_{σ_i} when $\sigma_i = \prod_{j=1}^I P_j$ for all i.

We shall heavily exploit the fact that v_{i,θ_i} is a linear transformation of π_{i,θ_i}. To see this, first look at the choice-specific expected return:

$$E\left[V_{i,\theta_i}(s_{t+1})\right| x_t, a_{it} = a_i] = E\left[E\left[V_{i,\theta_i}(s_{t+1})\right| x_{t+1}]\right| x_t, a_{it} = a_i], \text{ and}$$

$$E\left[V_{i,\theta_i}(s_t)\right| x_t] = E[\pi_{i,\theta_i}(a_{it}, a_{-it}, x_t) + \sum_{a'=0}^K \varepsilon_{it}(a')\mathbf{1}\left[a_{it} = a'\right]|x_t] + \beta E\left[E\left[V_{\theta_i}(s_{t+1})|x_{t+1}\right]|x_t\right].$$

Let $m_{i,\theta_i} = E[V_{i,\theta_i}(s_{it})|x_t = \cdot]$ and $g_{i,\theta_i} = E[V_{i,\theta_i}(s_{it+1})|x_t = \cdot, a_{it} = \cdot]$. Then, using a linear functional notation, we have

$$
\begin{aligned}
g_{i,\theta_i} &= \mathcal{H}_i m_{i,\theta_i}, \\
m_{i,\theta_i} &= r_{i,\theta_i} + \underline{r}_i + \mathcal{L} m_{i,\theta_i}, \quad \text{where for all } a, x \\
r_{i,\theta_i}(x) &= E\left[\pi_{i,\theta_i}(a_{it}, a_{-it}, x_t) \middle| x_t = x\right], \\
\underline{r}_i(x) &= E[\sum_{a'=0}^{K} \varepsilon_{it}(a') \mathbf{1}[a_{it} = a'] \, | x_t = x], \\
\mathcal{L} m(x) &= \beta E\left[m(x_{t+1}) | x_t = x\right], \\
\mathcal{H}_i m(a, x) &= E\left[m(x_{t+1}) | x_t = x, a_{it} = a\right],
\end{aligned}
$$

where $\mathcal{L}$ and $\mathcal{H}_i$ are linear maps and r_{i,θ_i} is a linear transformation of π_{i,θ_i}. Since $(I - \mathcal{L})^{-1}$ is also generally a well-defined linear map, as $\mathcal{L}$ is a contraction as its norm is strictly less than 1, then

$$v_{i,\theta_i} = \left(\mathcal{R}_i + \beta \mathcal{H}_i (I - \mathcal{L})^{-1} \mathcal{R}\right) \pi_{i,\theta_i} + \underline{v}_i,$$

where $\mathcal{R}_i$ and $\mathcal{R}$ are conditional expectation operators, conditioning on x_t, integrating over a_{-it} and a_t respectively , and $\underline{v}_i = \beta \mathcal{H}_i (I - \mathcal{L})^{-1} \underline{r}_i$.

The choice probabilities can also be written in terms of differences in choice specific expected payoffs. Let $\Delta v_{i,\theta_i}(a_i, x)$ denote $v_{i,\theta_i}(a_i, x) - v_{i,\theta_i}(0, x)$ for $a_i > 0$, then (3.2.4) becomes

$$P_{i,\theta_i}(a|x) = \Pr\left[\Delta v_{i,\theta_i}(a_i, x_t) + \varepsilon_{it}(a_i) > \Delta v_{i,\theta_i}(a_i', x_t) + \varepsilon_{it}(a_i') \text{ for all } a_i' > 0 | x_t = x\right]. \tag{3.2.5}$$

Since A and X are finite, the relationship between $\{\Delta v_{i,\theta_i}(a_i, x)\}_{a_i > 0, x \in X}$ and $\{\pi_{i,\theta_i}(\mathbf{a}, x)\}_{\mathbf{a} \in \mathbf{A}, x \in X}$ can be represented through a matrix equation. We state this representation as a lemma.

Lemma R: Under M1 - M4 $\{\Delta v_{i,\theta_i}(a_i, x)\}_{a_i > 0, x \in X}$ can then be represented by a $JK-$vector, $\Delta \mathbf{v}_{i,\theta_i}$:

$$\Delta \mathbf{v}_{i,\theta_i} = \mathbf{D}\left(\mathbf{R}_i + \beta \mathbf{H}_i \mathbf{M} \mathbf{R}\right) \pi_{i,\theta_i} + \Delta \underline{\mathbf{v}}_i, \tag{3.2.6}$$

where π_{i,θ_i} is a $J(K+1)^I-$vector of $\{\pi_{i,\theta_i}(a, x)\}_{a \in A, x \in X}$ so that elements in: $\mathbf{R}_i \pi_{i,\theta_i}$ are $\{E[\pi_{i,\theta_i}(a_i, a_{-it}, x_t)| x_t = x]\}_{a_i \in A, x \in X}$; $\mathbf{R}\pi_{i,\theta_i}$ are $\{E[\pi_{i,\theta_i}(a_{it}, a_{-it}, x_t)| x_t = x]\}_{x \in X}$; $\mathbf{M}$ involve
$\{\Pr[x_{t+1} = x' | x_t = x]\}$; $\mathbf{H}_i$ are $\Pr[x_{t+1} = x' | x_t = x, a_{it} = a_i]$; and, $\mathbf{D}$ is a difference matrix

with respect to the expected payoffs from playing action 0; and, $\Delta\underline{\mathbf{v}}_i$ is the differenced vector form of the transformation of $\underline{r}_i$ by $\beta_i\mathcal{H}_i(I-\mathcal{L})^{-1}$ normalized by action 0. The detailed constructions of $\Delta\underline{\mathbf{v}}_i, \mathbf{D}, \mathbf{R}_i, \mathbf{R}, \mathbf{H}_i$ and $\mathbf{M}$ are provided in the Appendix.

In what follows, we let $\Delta\mathbf{v}_i$ denote $\Delta\mathbf{v}_{i,\theta_{i0}}$. And, similarly, it shall be convenient to vectorize the probabilities. In particular, we let $\mathbf{P}_{i,\theta_i}$ and $\mathbf{P}_i$ denote the $JK-$vector that represent $\{P_{i,\theta_i}(a_i|x)\}_{a_i>0,x\in X}$ and $\{P_i(a_i|x)\}_{a_i>0,x\in X}$ respectively.

3.2.4 Estimation

Many objective functions proposed in the literature often can be written directly in terms of the probabilities from the pseudo-model, such as pseudo-likelihood and GMM, based on the construction that $\mathbf{P}_{i,\theta_i}$ coincides with $\mathbf{P}_i$ when $\theta_i=\theta_{i0}$. However, from a numerical perspective, computing the pseudo-probabilities requires a costly additional step of computation, namely the integration with respect to the distribution of ε_{it} that maps $\Delta\mathbf{v}_{i,\theta_i}$ into $\mathbf{P}_{i,\theta_i}$ (see (3.2.5)). These integrals generally do not have closed-form in the expected payoffs outside the well-known exception when private values are i.i.d. extreme value. Even if the integrals have closed-form, the integration is generally a nonlinear mapping of $\Delta\mathbf{v}_{i,\theta_i}$ into $\mathbf{P}_{i,\theta_i}$. In order to preserve the linear structure outlined previously, we propose to construct objective functions based directly on $\Delta\mathbf{v}_{i,\theta_i}$.

The validity of such objective functions, to identify θ_0, follows from the bijective relation between $\Delta\mathbf{v}_{i,\theta_i}$ and $\mathbf{P}_{i,\theta_i}$ for each i. This well-known result follows from Proposition 1 of Hotz and Miller (1993), which we shall refer to as Hotz and Miller's inversion in this paper (also see Lemma 8 of Matzkin (1991), Lemma 1 of Pesendorfer and Schmidt-Dengler (2008), and, for a recent generalization of these results, Norets and Takahashi (2013)).[5] In particular, it immediately follows that for any θ_i, $\mathbf{P}_{i,\theta_i}$ coincides with $\mathbf{P}_i$ if and only if $\Delta\mathbf{v}_{i,\theta_i}$ coincides with $\Delta\mathbf{v}_i$, where $\Delta\mathbf{v}_i$ is identifiable from the data by Hotz and Miller's inversion. Then we can construct a class of estimators based on minimizing the distance between $\{\Delta\mathbf{v}_{i,\theta_i}\}_{i=1}^I$ and $\{\Delta\mathbf{v}_i\}_{i=1}^I$.

Using Lemma R, we can write $\Delta\mathbf{v}_{i,\theta_i}=\mathcal{X}_i(\theta_i)+\Delta\underline{\mathbf{v}}_i$, where

$$\mathcal{X}_i(\theta_i)=\mathbf{D}(\mathbf{R}_i+\beta\mathbf{H}_i\mathbf{M}\mathbf{R})\,\pi_{i,\theta_i}. \tag{3.2.7}$$

[5] Pesendorfer and Schmidt-Dengler (2008) also show equilibrium condition can be characterized in terms of expected payoffs; see details of their Lemma 1 for further discussions.

Note that θ_i enters $\mathcal{X}_i(\theta_i)$ through a matrix transform of the vector π_{i,θ_i}, where the former does not depend on θ_i and the latter is completely known and specified by the researcher. By Hotz and Miller's inversion, we also have $\Delta\mathbf{v}_i = \Phi_i(\mathbf{P}_i)$ for some nonlinear, but known, function Φ_i that only depends on the distributional assumption of ε_{it}. Then we can define a $JK-$vector, $\mathcal{Y}_i$, where

$$\mathcal{Y}_i = \Phi_i(\mathbf{P}_i) - \Delta\underline{\mathbf{v}}_i. \tag{3.2.8}$$

Note that $\mathcal{Y}_i$ is defined independently of θ_i. So that, by construction:

$$\mathcal{Y}_i = \mathcal{X}_i(\theta_i) \quad \text{when} \quad \theta_i = \theta_{i0}.$$

Let $\mathcal{Y} = \left(\mathcal{Y}_1^\top, \ldots, \mathcal{Y}_I^\top\right)^\top$, $\theta = \left(\theta_1^\top, \ldots, \theta_I^\top\right)^\top$ and define a block diagonal matrix $\mathcal{X}(\theta) = diag(\mathcal{X}_1(\theta_1),$
$\ldots, \mathcal{X}_I(\theta_I))$. In the next section we analyze the asymptotic properties for a class of estimators that are motivated from minimizing

$$\mathcal{S}(\theta; \mathcal{W}) = (\mathcal{Y} - \mathcal{X}(\theta))^\top \mathcal{W}(\mathcal{Y} - \mathcal{X}(\theta)), \tag{3.2.9}$$

over Θ, for some weighting matrix $\mathcal{W}$.

It is also worth emphasizing that, through $\{\Delta\underline{\mathbf{v}}_i\}_{i=1}^I, \{\mathbf{R}_i\}_{i=1}^I, \mathbf{R}, \mathbf{L}$ and $\{\mathbf{H}_i\}_{i=1}^I$, for any θ: $\mathcal{X}(\theta)$ and $\mathcal{Y}$ are explicit functions, say $\mathcal{T}_{\mathcal{X}}(\theta; \gamma_0)$ and $\mathcal{T}_{\mathcal{Y}}(\gamma_0)$ respectively, of a finite-dimensional vector, γ_0, that consists of choice and transition probabilities. However, optimization with $\mathcal{S}(\theta; \mathcal{W})$ is infeasible since $\mathcal{X}(\theta)$ and $\mathcal{Y}$ are not observed, as γ_0 is unknown. Given a sample from a single equilibrium, $\{\alpha_1^*(s_{1t}), \ldots, \alpha_I^*(s_{It}), x_t, x_{t+1}\}$, γ_0 can be identified from the data under weak conditions, hence $\mathcal{X}(\theta)$ and $\mathcal{Y}$ can also be estimated directly from the data for all θ. Consequently we consider a feasible estimation criterion where $\mathcal{X}$ and $\mathcal{Y}$ are replaced by $\widehat{\mathcal{X}}(\theta) = \mathcal{T}_{\mathcal{X}}(\theta; \widehat{\gamma})$ and $\widehat{\mathcal{Y}} = \mathcal{T}_{\mathcal{Y}}(\widehat{\gamma})$ respectively, for some preliminary estimator, $\widehat{\gamma}$, of γ_0. We denote the sample counterpart of $\mathcal{S}$ by $\widehat{\mathcal{S}}$, so that

$$\widehat{\mathcal{S}}(\theta; \widehat{\mathcal{W}}) = (\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}(\theta))^\top \widehat{\mathcal{W}}(\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}(\theta)), \tag{3.2.10}$$

where $\widehat{\mathcal{W}}$ can be random and depend on the sample size. We define our estimator, $\widehat{\theta}(\widehat{\mathcal{W}})$, to be the minimizer of $\widehat{\mathcal{S}}\left(\theta;\widehat{\mathcal{W}}\right)$:

$$\widehat{\theta}(\widehat{\mathcal{W}}) = \arg\min_{\theta\in\Theta} \widehat{\mathcal{S}}(\theta;\widehat{\mathcal{W}}).$$

Therefore $\widehat{\theta}(\widehat{\mathcal{W}})$ is generally a nonlinear least square estimator with generated regressors and regressands. Note that $\widehat{\mathcal{S}}(\theta;\widehat{\mathcal{W}})$ is easy to evaluate for different values of θ, following (3.2.7) and (3.2.8), $\widehat{\mathcal{X}}_i(\theta)$ can be computed by a matrix multiplication of π_{i,θ_i} by the estimator of $\mathbf{D}(\mathbf{R}_i + \beta\mathbf{H}_i\mathbf{M}\mathbf{R})$, which does not depend on θ_i, and $\widehat{\mathcal{Y}}_i$ is also independent of θ_i.

3.3 Main Results

We give large sample properties of our estimator in full generality in Section 3.3.1. We consider special cases when payoffs have linear-in-parameter and partially linear specifications in 3.3.2 and 3.3.3 respectively. We discuss the relationship between our estimator and ALSE$_{PSD}$ in Section 3.3.4. In what follows we denote the matrix norm by $\|\cdot\|$, so that $\|B\| = \sqrt{trace\left(B^\top B\right)}$ for any real matrix B, and we let "$\xrightarrow{p}$" and "$\xrightarrow{d}$" denote convergence in probability and distribution respectively.

3.3.1 General Case

From the previous section, we see that $\mathcal{T}_{\mathcal{X}}(\theta;\cdot)$ and $\mathcal{T}_{\mathcal{Y}}(\cdot)$ are deterministic and smooth functions in γ for any θ. To analyze the asymptotic properties of $\widehat{\theta}(\widehat{\mathcal{W}})$, it will be useful to keep separate the sampling distribution of the preliminary estimator and the corresponding generated regressors and regressands. We begin with a preliminary requirement for $\widehat{\gamma}$.

Assumption P: *(i)* $\widehat{\gamma} \xrightarrow{p} \gamma_0$*; and, A(ii)* $\sqrt{N}\left(\widehat{\gamma}-\gamma_0\right) \xrightarrow{d} \mathcal{N}\left(0,\Xi\right)$.

There are several choices for $\widehat{\gamma}$ in practice that satisfy P under very weak conditions. The simplest options are perhaps the empirical choice and transition probabilities, otherwise kernel estimators can be employed (Li and Racine (2006)). We now present our regularity conditions and main results in terms of $(\mathcal{X}(\theta),\mathcal{Y})$ and their estimators $(\widehat{\mathcal{X}}(\theta),\widehat{\mathcal{Y}})$.

Assumption A1: $\theta_0 \in int(\Theta)$ *where* Θ *is a compact subset of* $\mathbb{R}^p$*, and* $\mathcal{X}(\theta) = \mathcal{X}(\theta_0)$ *if and only if* $\theta = \theta_0$.

Assumption A2: $\widehat{\mathcal{W}} \xrightarrow{p} \mathcal{W}$*, where* $\mathcal{W}$ *is a non-stochastic positive definite matrix.*

Assumption A3: $\sup_{\theta\in\Theta}\|\mathcal{X}(\theta)\|$ *and* $\|\mathcal{Y}\|$ *are finite, and* $\sup_{\theta\in\Theta}\left\|\widehat{\mathcal{X}}(\theta)-\mathcal{X}(\theta)\right\| \xrightarrow{p} 0$ *and* $\widehat{\mathcal{Y}} \xrightarrow{p} \mathcal{Y}$.

Assumption A4: $\mathcal{X}(\theta)$ *is continuously differentiable at* θ_0 *and* $\nabla_{\mathcal{X}} = \left.\frac{\partial\mathcal{X}(\theta)}{\partial\theta^\top}\right|_{\theta=\theta_0}$ *has full column rank.*

Assumption A5: $\sup_{\theta\in B_\delta(\theta_0)}\left\|\frac{\partial\widehat{\mathcal{X}}(\theta)}{\partial\theta^\top}-\frac{\partial\mathcal{X}(\theta)}{\partial\theta^\top}\right\| \xrightarrow{p} 0$*, where* $B_\delta(\theta_0)$ *denotes some neighborhood of* θ_0.

Define $\widehat{\mathcal{U}} = \widehat{\mathcal{Y}} - \widehat{\mathcal{X}}(\theta_0)$.

Assumption A6: $\sqrt{N}\widehat{\mathcal{U}} \xrightarrow{d} \mathcal{N}(0,\Sigma)$ *for some non-stochastic positive definite matrix* Σ.

Comments on Assumptions A1 - A3.

These conditions are sufficient for the consistency of our estimator. A1 - A2 constitute to a high level identification condition as it ensures (3.2.9) has a unique solution at θ_0. There has been little work on more primitive conditions for parametric identification of payoff functions in dynamic games. Most identification results in the literature are nonparametric that build on the work of Magnac and Thesmar (2002); see Pesendorfer and Schmidt-Dengler (2008) and Bajari et al. (2009). However, using Hotz and Miller's inversion, it follows that the condition for identification of the pseudo-model at θ_0, in the sense that $P_{i,\theta_i} = P_{i,\theta_{i0}}$ for all i if and only if $\theta_i = \theta_{i0}$ for all i, is precisely the identification condition required in A1. Furthermore, by inspecting Lemma R more closely, for each i, we see that the necessary and sufficient condition for the unique parameterization of $\mathcal{X}_i(\theta_i)$ at θ_{i0} is for the intersection between the $\{\pi_{i,\theta_i} - \pi_{i,\theta_{i0}} : \theta_i \in \Theta_i\backslash\{\theta_{i0}\}\}$ and the null space of $\mathbf{D}(\mathbf{R}_i + \beta\mathbf{H}_i\mathbf{M}\mathbf{R})$ to be empty. Although, without any restriction on π_{i,θ_i}, A1 generally does not hold since $\mathbf{D}(\mathbf{R}_i + \beta\mathbf{H}_i\mathbf{M}\mathbf{R})$ is always rank-deficient. For a closely related discussion see Srisuma (2013), who provides constructive conditions for parametric identification results in a single agent model that can be generalized directly to the games considered in this paper. Also see the identification condition and comments of B1 in Section 3.3.2 when linear-in-parameter restriction is imposed. The uniform boundedness and consistency conditions essentially depend on $\{\pi_{i,\theta_i}\}_{i=1}^I$. In particular, if $\mathbf{D}(\mathbf{R}_i + \beta\mathbf{H}_i\mathbf{M}\mathbf{R})$ is finite then continuity of $\{\pi_{i,\theta_i}\}_{i=1}^I$

ensures $\sup_{\theta\in\Theta} \|\mathcal{X}(\theta)\|$ is finite since Θ is compact. Then uniform consistency also follows if there exists a consistent estimator for $\mathbf{D}(\mathbf{R}_i + \beta\mathbf{H}_i\mathbf{M}\mathbf{R})$, which is implied by P(i).

Comments on Assumptions A4 - A6.

For the distribution theory, additional local conditions around θ_0 are required. A4 - A5 are standard smoothness and regularity conditions for an asymptotic normality of an extremum estimator that optimizes a smooth objective function. Similar to the discussion of sufficient conditions for A3, using Lemma R, a sufficient condition for continuous differentiability of $\mathcal{X}(\theta)$ in A4 is continuous differentiability of π_{i,θ_i} at θ_{i0} for all i, then A5 will also follow if P(i) holds. Furthermore, if P(ii) holds, so that the elements in $\widehat{\mathcal{X}}(\theta_0)$ and $\widehat{\mathcal{Y}}$ have asymptotically normal distribution, then by applying a delta-method A6 also holds with $\Sigma = \nabla_\gamma \Xi \nabla_\gamma^\top$, where $\nabla_\gamma = \frac{\partial}{\partial\gamma^\top}\left(\mathcal{T}_{\mathcal{Y}}(\gamma) - \mathcal{T}_{\mathcal{X}}(\theta_0;\gamma)\right)|_{\gamma=\gamma_0}$.

Our estimators are consistent and asymptotically normal under these assumptions.

Theorem 1 (Consistency): *Under assumptions A1 - A3,* $\widehat{\theta}(\widehat{\mathcal{W}}) \xrightarrow{p} \theta_0$.

Theorem 2 (Asymptotic Normality): *Under assumptions A1 - A6,*

$$\sqrt{N}\left(\widehat{\theta}(\widehat{\mathcal{W}}) - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, \Omega_{\mathcal{W}}\right),$$

where $\Omega_{\mathcal{W}} = \left(\nabla_{\mathcal{X}}^\top \mathcal{W} \nabla_{\mathcal{X}}\right)^{-1} \nabla_{\mathcal{X}}^\top \mathcal{W}\Sigma\mathcal{W}\nabla_{\mathcal{X}} \left(\nabla_{\mathcal{X}}^\top \mathcal{W} \nabla_{\mathcal{X}}\right)^{-1}$.

In large sample, the estimators that uniquely solve (3.2.10) are distinguishable up to the first order by $\Omega_{\mathcal{W}}$. The efficient estimator in this class can be found by choosing the optimal weighting matrix, $\mathcal{W}^*$, that minimizes $\Omega_{\mathcal{W}}$ over the set of all possible positive definite matrices (i.e. efficiency gain in the spirit of Chamberlain (1982) and Hansen (1982) for instance).

Theorem 3 (Efficient Estimator): *Under assumptions A1 - A6, (i) the asymptotic variance of* $\sqrt{N}\left(\widehat{\theta}(\widehat{\mathcal{W}}) - \theta_0\right)$ *is bounded below by* $\Omega_{\Sigma^{-1}} = \left(\nabla_{\mathcal{X}}^\top \Sigma^{-1} \nabla_{\mathcal{X}}\right)^{-1}$*; and, (ii) if* $\widehat{\mathcal{W}} \xrightarrow{p} \Sigma^{-1}$ *then* $\sqrt{N}\left(\widehat{\theta}(\widehat{\mathcal{W}}) - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, \Omega_{\Sigma^{-1}}\right)$.

The first part of Theorem 3 says that the lower variance bound for the class of estimators we consider is $\left(\nabla_{\mathcal{X}}^\top \Sigma^{-1} \nabla_{\mathcal{X}}\right)^{-1}$. The second part states that any consistent estimator of Σ^{-1} is sufficient to produce an efficient estimator. In practice, consistent estimator for Σ^{-1} will typically require a preliminary consistent estimator for θ_0. The simplest choice is to choose $\mathcal{W}$ to be an identity matrix, I_d. In this case the estimator for θ_{i0} can be computed individually for each player. We state this in the following corollary.

Corollary A (Identity Weighted Estimator): *Under assumptions A1, A3 - A6,* $\sqrt{N}\left(\widehat{\theta}(I_d)-\theta_0\right) \xrightarrow{d} \mathcal{N}\left(0,\Omega_{I_d}\right)$*, where* $\widehat{\theta}(I_d) = \left(\widehat{\theta}_1(I_d)^\top,\ldots,\widehat{\theta}_I(I_d)^\top\right)^\top$*. Furthermore, for all* i*:* $\widehat{\theta}_i(I_d) = \arg\min_{\theta_i\in\Theta_i}(\widehat{\mathcal{Y}}_i - \widehat{\mathcal{X}}_i\left(\theta_i\right))^\top(\widehat{\mathcal{Y}}_i - \widehat{\mathcal{X}}_i\left(\theta_i\right))$ *such that* $\sqrt{N}\left(\widehat{\theta}_i(I_d)-\theta_{i0}\right) \xrightarrow{d}$ $\mathcal{N}(0,\left(\nabla_{\mathcal{X}_i}^\top\nabla_{\mathcal{X}_i}\right)^{-1}$
$\nabla_{\mathcal{X}_i}^\top\Sigma_i\nabla_{\mathcal{X}_i}\left(\nabla_{\mathcal{X}_i}^\top\nabla_{\mathcal{X}_i}\right)^{-1})$ *with* $\nabla_{\mathcal{X}_i} = \left.\frac{\partial\mathcal{X}_i(\theta)}{\partial\theta_i^\top}\right|_{\theta_i=\theta_{i0}}$ *and* $\Sigma_i = \lim_{N\to\infty} Var(\sqrt{N}(\widehat{\mathcal{Y}}_i - \mathcal{Y}_i - (\widehat{\mathcal{X}}_i\left(\theta_{i0}\right) - \mathcal{X}_i\left(\theta_{i0}\right)))$*.*

3.3.2 Linear-in-Parameter Specification

We now consider the leading special case when payoff functions have a linear-in-parameter specification.

Assumption M5 *(Linear-in-parameter payoffs). For all* $(i,\theta_i,a_i,a_{-i},x)$*,*

$$\pi_{i,\theta_i}\left(a_i,a_{-i},x\right) = \theta_i^\top\pi_i\left(a_i,a_{-i},x\right),$$

for some $p-$*dimensional vector* $\pi_i\left(a_i,a_{-i},x\right) = \left(\pi_i^1\left(a_i,a_{-i},x\right),\ldots,\pi_i^p\left(a_i,a_{-i},x\right)\right)^\top$*.*

We assume M1 - M5 hold throughout this subsection. Then, with a slight abuse of notation, $\mathcal{X}_i\left(\theta_i\right)$ in (3.2.7) simplifies to $\mathcal{X}_i\theta_i$, where

$$\mathcal{X}_i = \mathbf{D}\left(\mathbf{R}_i + \beta\mathbf{H}_i\mathbf{M}\mathbf{R}\right)\mathbf{\Pi}_i, \tag{3.3.1}$$

and $\mathbf{\Pi}_i$ is a $J\left(K+1\right)^I$ by p matrix of $\left\{\pi_i\left(a_i,a_{-i},x\right)\right\}_{a_i\in A,x\in X}$. Let $\mathcal{X} = diag\left(\mathcal{X}_1,\ldots,\mathcal{X}_I\right)$. The limiting and sample objective functions defined in (3.2.9) and (3.2.10) respectively become

$$\begin{aligned}\mathcal{S}^{lip}\left(\theta;\mathcal{W}\right) &= (\mathcal{Y}-\mathcal{X}\theta)^\top\mathcal{W}(\mathcal{Y}-\mathcal{X}\theta), \text{ and}\\ \widehat{\mathcal{S}}^{lip}(\theta;\widehat{\mathcal{W}}) &= (\widehat{\mathcal{Y}}-\widehat{\mathcal{X}}\theta)^\top\widehat{\mathcal{W}}(\widehat{\mathcal{Y}}-\widehat{\mathcal{X}}\theta).\end{aligned}$$

If $\widehat{\mathcal{X}}^\top\widehat{\mathcal{W}}\widehat{\mathcal{X}}$ is non-singular, then $\widehat{\mathcal{S}}^{lip}(\theta;\widehat{\mathcal{W}})$ is globally convex. The solution to the minization problem has a well-known closed-form expression of a weighted least squares estimator, namely

$$\widehat{\theta}^{lip}(\widehat{\mathcal{W}}) = \left(\widehat{\mathcal{X}}^{\top}\widehat{\mathcal{W}}\widehat{\mathcal{X}}\right)^{-1}\widehat{\mathcal{X}}^{\top}\widehat{\mathcal{W}}\widehat{\mathcal{Y}}. \tag{3.3.2}$$

Although the large sample properties for $\widehat{\theta}^{lip}(\widehat{\mathcal{W}})$ follow immediately from Section 3.3.1, they can be specialized substantially to incorporate M5. Since the results in this subsection may be most relevant for empirical applications we provide some details here.

Assumption B1: *$\mathcal{X}$ has full column rank.*

Assumption B2: *$\widehat{\mathcal{W}} \xrightarrow{p} \mathcal{W}$, where $\mathcal{W}$ is a non-stochastic positive definite matrix.*

Assumption B3: *$\|\mathcal{X}\|$ and $\|\mathcal{Y}\|$ are finite, and $\widehat{\mathcal{X}} \xrightarrow{p} \mathcal{X}$ and $\widehat{\mathcal{Y}} \xrightarrow{p} \mathcal{Y}$.*

Define $\widehat{\mathcal{U}}^{lip} = \widehat{\mathcal{Y}} - \widehat{\mathcal{X}}\theta_0$.

Assumption B4: *$\sqrt{N}\widehat{\mathcal{U}}^{lip} \xrightarrow{d} \mathcal{N}\left(0, \Sigma^{lip}\right)$ for some non-stochastic positive definite matrix Σ^{lip}.*

Comments on Assumptions B1 - B4.

Similar to A1 - A2, B1 and B2 ensure $\mathcal{S}^{lip}(\theta; \mathcal{W})$ has a unique solution at θ_0. In this case, the full rank condition of $\mathcal{X}$ is a necessary and sufficient condition for the identification of the pseudo-model (for more details see Srisuma (2013)). The sample counterpart of B1, namely the rank condition of $\widehat{\mathcal{X}}$, also has a finite sample significance. If $\widehat{\mathcal{W}}$ is positive definite, then the full column rank condition of $\widehat{\mathcal{X}}$ is necessary and sufficient for $\widehat{\mathcal{S}}^{lip}(\theta; \widehat{\mathcal{W}})$ to have a unique solution, which equals to $\widehat{\theta}^{lip}(\widehat{\mathcal{W}})$ as defined in (3.3.2). Assumptions B3 and B4 are immediate specializations of A3 - A6.

We state the large sample properties for $\widehat{\theta}^{lip}(\widehat{\mathcal{W}})$ as corollaries without proofs.

Corollary 1 (Consistency): *Under assumptions B1 - B3, $\widehat{\theta}^{lip}(\widehat{\mathcal{W}}) \xrightarrow{p} \theta_0$.*

Corollary 2 (Asymptotic Normality): *Under assumptions B1 - B4,*

$$\sqrt{N}\left(\widehat{\theta}^{lip}(\widehat{\mathcal{W}}) - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, \Omega^{lip}_{\mathcal{W}}\right),$$

where $\Omega^{lip}_{\mathcal{W}} = \left(\mathcal{X}^{\top}\mathcal{W}\mathcal{X}\right)^{-1}\mathcal{X}^{\top}\mathcal{W}\Sigma^{lip}\mathcal{W}\mathcal{X}\left(\mathcal{X}^{\top}\mathcal{W}\mathcal{X}\right)^{-1}$.

Corollary 3 (Efficient Estimator): *Under assumptions B1 -B4, (i) the asymptotic variance of* $\sqrt{N}\left(\widehat{\theta}^{lip}(\widehat{\mathcal{W}})-\theta_0\right)$ *is bounded below by* $\Omega^{lip}_{\Sigma^{lip^{-1}}}=\left(\mathcal{X}^\top\Sigma^{lip^{-1}}\mathcal{X}\right)^{-1}$*; and, (ii) if* $\widehat{\mathcal{W}}\xrightarrow{p}\Sigma^{lip^{-1}}$ *then* $\sqrt{N}\left(\widehat{\theta}^{lip}(\widehat{\mathcal{W}})-\theta_0\right)\xrightarrow{d}\mathcal{N}\left(0,\Omega^{lip}_{\Sigma^{lip^{-1}}}\right)$.

Similarly to the general case, consistent estimator for $\Sigma^{lip^{-1}}$ requires a preliminary consistent estimator for θ_0. We have the counterpart to Corollary A when we choose $\mathcal{W}$ to be an identity matrix $\mathbf{I}$.

Corollary B (Identity Weighted Estimator): *Under assumptions B1, B3 and B4,* $\sqrt{N}\left(\widehat{\theta}^{lip}(\mathbf{I})-\theta_0\right)\xrightarrow{d}\mathcal{N}\left(0,\Omega^{lip}_{\mathbf{I}}\right)$*, where* $\widehat{\theta}^{lip}(\mathbf{I})=\left(\widehat{\mathcal{X}}^\top\widehat{\mathcal{X}}\right)^{-1}\widehat{\mathcal{X}}^\top\widehat{\mathcal{Y}}$ *and*
$\Omega^{lip}_{\mathbf{I}}=\left(\mathcal{X}^\top\mathcal{X}\right)^{-1}\mathcal{X}^\top\Sigma\mathcal{X}\left(\mathcal{X}^\top\mathcal{X}\right)^{-1}$*. Furthermore, for all i:* $\widehat{\theta}^{lip}(\mathbf{I})=\left(\widehat{\theta}^{lip}_1(\mathbf{I})^\top,\ldots,\widehat{\theta}^{lip}_I(\mathbf{I})^\top\right)^\top$ *such that* $\widehat{\theta}^{lip}_i(\mathbf{I})=\left(\widehat{\mathcal{X}}_i^\top\widehat{\mathcal{X}}_i\right)^{-1}\widehat{\mathcal{X}}_i^\top\widehat{\mathcal{Y}}_i$ *and* $\sqrt{N}\left(\widehat{\theta}^{lip}_i(\mathbf{I})-\theta_{i0}\right)\xrightarrow{d}\mathcal{N}(0,\left(\mathcal{X}_i^\top\mathcal{X}_i\right)^{-1}\mathcal{X}_i^\top\Sigma^{lip}_i\mathcal{X}_i\left(\mathcal{X}_i^\top\mathcal{X}_i\right)^{-1})$ *with* $\Sigma^{lip}_i=\lim_{N\to\infty}Var(\sqrt{N}(\widehat{\mathcal{Y}}_i-\mathcal{Y}_i-(\widehat{\mathcal{X}}_i-\mathcal{X}_i)\theta_{i0})$.

We have shown here that once we have $(\widehat{\mathcal{Y}},\widehat{\mathcal{X}})$, under some regularity conditions, a consistent estimator for θ_0 can be obtained by an OLS estimator, $\widehat{\theta}^{lip}(\mathbf{I})=\left(\widehat{\mathcal{X}}^\top\widehat{\mathcal{X}}\right)^{-1}\widehat{\mathcal{X}}^\top\widehat{\mathcal{Y}}$ (Corollary B), which can be used to construct an efficient estimator using a familiar a feasible GLS formulation, $\widehat{\theta}^{lip}\left(\widehat{\Sigma}^{lip^{-1}}\right)=\left(\widehat{\mathcal{X}}^\top\widehat{\Sigma}^{lip^{-1}}\widehat{\mathcal{X}}\right)^{-1}\widehat{\mathcal{X}}^\top\widehat{\Sigma}^{lip^{-1}}\widehat{\mathcal{Y}}$ where $\widehat{\Sigma}^{lip^{-1}}$ is a consistent estimator of $\Sigma^{lip^{-1}}$.

Our closed-form estimators also readily accommodate linear restrictions. For instance, sometimes there are a priori restrictions one may wish to impose on θ_0 such as symmetry. More formally, suppose θ_0 is known to satisfy $\mathcal{D}^\top\theta_0=\delta$ for some known p by q matrix $\mathcal{D}$ that has full row rank $q<p$ and some $q-$dimensional vector δ. Then a restricted estimator $\widetilde{\theta}^{lip}(\widehat{\mathcal{W}})$ that minimizes (3.2.10) subject to $\mathcal{D}^\top\widetilde{\theta}^{lip}(\widehat{\mathcal{W}})=\delta$, has the following closed-form expression

$$\widetilde{\theta}^{lip}(\widehat{\mathcal{W}})=\widehat{\theta}^{lip}(\widehat{\mathcal{W}})-\left(\widehat{\mathcal{X}}^\top\widehat{\mathcal{W}}\widehat{\mathcal{X}}\right)^{-1}\mathcal{D}\left(\mathcal{D}^\top\left(\widehat{\mathcal{X}}^\top\widehat{\mathcal{W}}\widehat{\mathcal{X}}\right)^{-1}\mathcal{D}\right)^{-1}\left(\mathcal{D}^\top\widehat{\theta}^{lip}(\widehat{\mathcal{W}})-\delta\right),$$

where $\widehat{\theta}^{lip}(\widehat{\mathcal{W}})$ is the unrestricted estimator defined in (3.3.2). The expression above can be derived using Lagrangean method or through matrix manipulations (see Amemiya (1985, Section 1.4)). And, since $\widetilde{\theta}^{lip}(\widehat{\mathcal{W}})$ is an affine transformation of $\widehat{\theta}^{lip}(\widehat{\mathcal{W}})$, it is easy to verify that the optimal weighting matrices for $\widetilde{\theta}^{lip}(\widehat{\mathcal{W}})$ are the same as those described in Corollary 3, i.e. any $\widehat{\mathcal{W}}\xrightarrow{p}\Sigma^{lip^{-1}}$.

3.3.3 Partially Linear Specification

One may argue that, in some situations, Assumption M5 is at odds with the spirit of structural estimation if the functions in the vector π_i are interpreted as basis functions. We relax the linear-in-parameter requirement and instead consider a partially linear structure that may arise naturally by ways of additive fixed effects, or, frequently in modeling of entry/exit games, as fixed costs or scrap value. Now suppose $\theta_i = \left(\theta_i^{A\top}, \theta_i^{B\top}\right)^\top$ for all i.

Assumption M6 *(Partially linear payoffs). For all $(i, \theta_i, a_i, a_{-i}, x)$,*

$$\pi_{i,\theta_i}(a_i, a_{-i}, x) = \theta_i^{A\top} \pi_i^A (a_i, a_{-i}, x) + \pi_{i,\theta_i^B}^B (a_i, a_{-i}, x),$$

for some $p-$dimensional vector $\pi_i^A (a_i, a_{-i}, x) = \left(\pi_i^{A1} (a_i, a_{-i}, x), \ldots, \pi_i^{Ap} (a_i, a_{-i}, x)\right)^\top$.

We assume M1 - M4 and M6 hold throughout this subsection. Then it is easy to see that the RHS of equation (3.2.6) in Lemma R becomes

$$\mathbf{D}(\mathbf{R}_i + \beta \mathbf{H}_i \mathbf{M}\mathbf{R}) \pi_{i,\theta_i}^A + \mathbf{D}(\mathbf{R}_i + \beta \mathbf{H}_i \mathbf{M}\mathbf{R}) \pi_{i,\theta_i^B}^B + \Delta \underline{\mathbf{v}}_i,$$

and, we define, analogously to (3.2.7) and (3.3.1), $\mathcal{X}_i^A = \mathbf{D}(\mathbf{R}_i + \beta \mathbf{H}_i \mathbf{M}\mathbf{R}) \mathbf{\Pi}_i^A$, and $\mathcal{X}_i^B \left(\theta_i^B\right) = \mathbf{D}(\mathbf{R}_i + \beta \mathbf{H}_i \mathbf{M}\mathbf{R}) \pi_{i,\theta_i^B}^B$. Once again, stacking up the vectors from all players, the limiting and sample objective functions defined in (3.2.9) and (3.2.10) respectively become

$$\begin{aligned} \mathcal{S}^{pl}(\theta; \mathcal{W}) &= (\mathcal{Y} - \mathcal{X}^A \theta^A - \mathcal{X}^B \left(\theta^B\right))^\top \mathcal{W} (\mathcal{Y} - \mathcal{X}^A \theta^A - \mathcal{X}^B \left(\theta^B\right)), \text{ and} \\ \widehat{\mathcal{S}}^{pl}(\theta; \widehat{\mathcal{W}}) &= (\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}^A \theta^A - \widehat{\mathcal{X}}^B \left(\theta^B\right))^\top \widehat{\mathcal{W}} (\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}^A \theta^A - \widehat{\mathcal{X}}^B \left(\theta^B\right)), \end{aligned}$$

where the terms in the above display should by now be familiar. In order to avoid repetition we only provide a brief discussion of how θ can be (efficiently) estimated.

The structural identifying condition in this setting is:

$$\mathcal{Y} = \mathcal{X}^A \theta^A + \mathcal{X}^B \left(\theta^B\right) \text{ if and only if } \left(\theta^A, \theta^B\right) = \left(\theta_0^A, \theta_0^B\right).$$

The additively linear structure allows us to use a Frisch-Waugh-Lovell type argument to estimate θ_0^A and θ_0^B sequentially in two stages. In particular, θ_0^A and θ_0^B satisfy the following identities:

$$\mathcal{M}_{\mathcal{W}A}\mathcal{Y} = \mathcal{M}_{\mathcal{W}A}\mathcal{X}^B\left(\theta_0^B\right), \tag{3.3.3}$$

where $\mathcal{M}_{\mathcal{W}A} = I - \mathcal{X}^A\left(\mathcal{X}^{A\top}\mathcal{W}\mathcal{X}^A\right)^{-1}\mathcal{X}^{A\top}\mathcal{W}$ is an oblique projection matrix (e.g. see Davidson and MacKinnon (1993)), so that $\mathcal{M}_{\mathcal{W}A}\mathcal{X}^A$ is a matrix of zeros, and

$$\mathcal{Y} - \mathcal{X}^B\left(\theta_0^B\right) = \mathcal{X}^A\theta_0^A. \tag{3.3.4}$$

An asymptotic least squares estimator that minimizes $\widehat{\mathcal{S}}^{pl}(\theta;\widehat{\mathcal{W}})$ can then be constructed sequentially in two stages. Let

$$\widehat{\mathcal{S}}_1^{pl}(\theta^B;\widehat{\mathcal{W}}) = (\mathcal{M}_{\widehat{\mathcal{W}}A}\widehat{\mathcal{Y}} - \mathcal{M}_{\widehat{\mathcal{W}}A}\widehat{\mathcal{X}}\left(\theta^B\right))^\top\widehat{\mathcal{W}}(\mathcal{M}_{\widehat{\mathcal{W}}A}\widehat{\mathcal{Y}} - \mathcal{M}_{\widehat{\mathcal{W}}A}\widehat{\mathcal{X}}\left(\theta^B\right)),$$

where $\mathcal{M}_{\widehat{\mathcal{W}}A} = I - \mathcal{X}^A(\mathcal{X}^{A\top}\widehat{\mathcal{W}}\mathcal{X}^A)^{-1}\mathcal{X}^{A\top}\widehat{\mathcal{W}}$. In the first stage we obtain $\widehat{\theta}^{plB}(\widehat{\mathcal{W}}) = \arg\min_{\theta^B}\widehat{\mathcal{S}}_1^{pl}(\theta^B;\widehat{\mathcal{W}})$. For the second stage, let

$$\widehat{\mathcal{S}}_2^{pl}(\theta^A;\widehat{\mathcal{W}}) = (\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}^B(\widehat{\theta}^B) - \widehat{\mathcal{X}}^A\theta^A)^\top\widehat{\mathcal{W}}(\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}^B(\widehat{\theta}^B) - \widehat{\mathcal{X}}^A\theta^A).$$

Then $\widehat{\theta}^{plA}(\widehat{\mathcal{W}}) = \arg\min_{\theta^A}\widehat{\mathcal{S}}_2^{pl}(\theta^A;\widehat{\mathcal{W}}) = (\widehat{\mathcal{X}}^{A\top}\widehat{\mathcal{W}}\widehat{\mathcal{X}}^A)^{-1}\widehat{\mathcal{X}}^{A\top}\widehat{\mathcal{W}}(\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}^B(\widehat{\theta}^B))$. It is easy to verify the first order conditions that $\widehat{\theta}^{plA}(\widehat{\mathcal{W}})$ and $\widehat{\theta}^{plB}(\widehat{\mathcal{W}})$ individually solve are identical to the ones obtained from jointly minimizing $\widehat{\mathcal{S}}^{pl}(\theta;\widehat{\mathcal{W}})$.

The practical advantage of the sequential approach is purely numerical, in the same spirit as the well-known partition regression methods described since the work of Frisch and Waugh (1933). Specifically, we only need to perform nonlinear optimization routine to search over a reduced parameter space for $\widehat{\theta}^{plB}(\widehat{\mathcal{W}})$ in the first stage, as $\widehat{\theta}^{plA}(\widehat{\mathcal{W}})$ has a closed-form expression in terms of $\widehat{\theta}^{plB}(\widehat{\mathcal{W}})$. Note also that the optimal weighting matrix for $\widehat{\mathcal{S}}_1^{pl}$ and $\widehat{\mathcal{S}}_2^{pl}$ is the same, and is identical to the one described in Theorem 3.

3.3.4 An Equivalent ALSE

Generally it is not possible to directly compare asymptotic efficiency of different estimators in the literature, although they estimate the same model, since many of the estimators are defined using non-nesting objective functions. An exception can be found in Pesendorfer and Schmidt-Dengler (2008), who show ALSE$_{PSD}$ includes some estimators of Aguirregabiria and Mira (2007) and Pakes, Ostrovsky and Berry (2007) as special cases. Similar to our general estimator defined in Section 3.2, the class of ALSE$_{PSD}$ is also indexed by a positive definite matrix and optimal weights can be found to define an efficient estimator (cf. Theorem 3). As implied by the Proposition E below, our efficient estimator is asymptotically equivalent to the efficient ALSE$_{PSD}$. In fact, more is true, the class of estimators we consider and that of Pesendorfer and Schmidt-Dengler are asymptotically equivalent in the sense that one can choose appropriate weighting matrices so that the two estimators always have the same asymptotic distribution.

Proposition E. *ALSE$_{PSD}$ and our estimator are asymptotically equivalent.*

The equivalence follows from the existence of a smooth bijective relation between the choice probabilities and the normalized expected payoffs, i.e. essentially by Hotz and Miller's inversion and an application of the inverse function theorem. The precise relationship between the two estimators are summarized by the equations in display (3.5.3) that can be found in the Appendix.

We end this section with a remark on the relationship between asymptotic least squares estimators and GMM estimators. ALSE$_{PSD}$ and our estimator are defined using objective functions that look at the differences between the data and pseudo-model implied probabilities and payoffs respectively at every possible actions and observed states. These differences can also be written as moment conditions, thus asymptotic least squares estimators can also equivalently be defined as GMM estimators (see Chamberlain (1987)). As a consequence, it follows from Proposition E that the GMM estimators of Hotz and Miller (1993) and Hotz et al. (1994) are also asymptotically equivalent for a stationary single agent decision model (a special case of our game when $I = 1$).[6]

3.4 Monte Carlo Experiments

We illustrate the performance of our closed-form estimator using the Monte Carlo design in Section 7 of Pesendorfer and Schmidt-Dengler (2008); who also provide further comparison with other estimators in the literature.

Setup

[6] The estimator of Hotz et al. (1994) has an additional source of sampling error since they estimate the discounted expected payoffs, $E\left[V_{i,\theta_i}(s_{t+1})\right| x_t, a_{it}]$, by forward simulation. However, under suitable conditions, the error from forward simulation does not affect the asymptotic distribution of their estimator.

Consider a symmetric two-firm dynamic entry game. In each period t, each firm $i(=1,2)$ has two possible choices: be active or not active, $a_{it} \in \{0,1\}$, where 0 corresponds to "not active" and 1 to "active" . Publically observed state variable has four elements, and can be represented by the actions made by both firms in period $t-1$, so that $x_t = (a_{1t-1}, a_{2t-1})$. The vector of states evolves over time according to the transition $s_{t+1} = a_t$. Firm $1'$s period payoffs are described as follows:

$$\pi_{1,\theta}(a_{1t}, a_{2t}, x_t) = \mathbf{1}[a_{1t}=1] \cdot [\theta_1 + \theta_2 a_{2t}] + \mathbf{1}[a_{1t}=1, a_{1t-1}=0] \cdot F + \mathbf{1}[a_{1t}=0, a_{1t-1}=1] \cdot W,$$

where $(\theta_1, \theta_2, F, W)$ denote respectively the monopoly profit, duopoly profit, entry cost and scrap value that firm 1 may obtain. Each firm also receives additive private shocks that are i.i.d. $\mathcal{N}(0,1)$. The game is symmetric and firm's 2 payoffs are defined analogously.

We set $(\theta_{10}, \theta_{20}, F_0, W_0) = (1.2, -1.2, -0.2, 0.1)$. Pesendorfer and Schmidt-Dengler (2008, p.920) show that there are three distinct equilibria (five if we permute the identity of the players as there is one symmetric equilibrium). We generate the data using different equilibria of the game and provide estimates for $(\theta_{10}, \theta_{20}, F_0)$ for each equilibrium. W is taken as known, since it is not separately identified (see Aguirregabiria and Suzuki (2013)). For each sample size $T = 100, 500, 1000, 5000$, we report the same statistics as Pesendorfer and Schmidt-Dengler (mean and standard deviation of each estimator, and average mean squared error) from 1000 simulations of four estimators: OLS, GLS, PSD-I and PSD-E, for each equilibrium. OLS and GLS estimators correspond to our inefficient and efficient estimators that have closed-form (see Corollary B and Corollary 3 respectively). PSD-I and PSD-E are the inefficient and efficient versions of ALSE_{PSD} respectively; the former uses identity weighting matrix. Our Tables 1 - 3 below correspond respectively to equilibria 1 - 3 in Pesendorfer and Schmidt-Dengler (2008), thus are directly comparable to their Tables 1 - 3 on p.921-922.

Table 3.1: Monte Carlo results (Equilibrium 1)

T	Estimator	F		θ_{10}		θ_{20}		MSE
100	OLS	-0.244	(0.328)	1.071	(0.33)	-1.087	(0.385)	0.396
	GLS	-0.210	(0.136)	1.227	(0.276)	-1.23	(0.255)	0.161
	PSD-I	-0.262	(0.316)	1.083	(0.341)	-1.094	(0.390)	0.395
	PSD-E	-0.175	(0.155)	1.292	(0.303)	-1.327	(0.301)	0.231
500	OLS	-0.213	(0.151)	1.169	(0.141)	-1.161	(0.179)	0.077
	GLS	-0.197	(0.048)	1.213	(0.133)	-1.209	(0.096)	0.029
	PSD-I	-0.220	(0.148)	1.176	(0.144)	-1.167	(0.186)	0.079
	PSD-E	-0.188	(0.047)	1.232	(0.129)	-1.223	(0.102)	0.031
1000	OLS	-0.206	(0.105)	1.184	(0.09)	-1.182	(0.125)	0.035
	GLS	-0.200	(0.030)	1.200	(0.081)	-1.197	(0.062)	0.011
	PSD-I	-0.209	(0.102)	1.186	(0.090)	-1.185	(0.130)	0.036
	PSD-E	-0.195	(0.029)	1.212	(0.077)	-1.204	(0.064)	0.011
5000	OLS	-0.204	(0.079)	1.194	(0.061)	-1.190	(0.093)	0.019
	GLS	-0.206	(0.074)	1.196	(0.059)	-1.192	(0.089)	0.017
	PSD-I	-0.201	(0.079)	1.199	(0.064)	-1.196	(0.094)	0.019
	PSD-E	-0.203	(0.077)	1.198	(0.061)	-1.195	(0.092)	0.018

Notes: OLS and GLS are our closed-form estimators that are inefficient and efficient respectively. PSD-I and PSD-E are asymptotic least squares estimators of Pesendorfer and Schmidt-Dengler (2008) that are inefficient (identity weighted) and efficient respectively.

Table 3.2: Monte Carlo results (Equilibrium 2)

T	Estimator	F		θ_{10}		θ_{20}		MSE
100	OLS	-0.317	(0.472)	0.971	(0.38)	-0.891	(0.543)	0.822
	GLS	-0.428	(0.333)	0.998	(0.328)	-0.892	(0.438)	0.598
	PSD-I	-0.264	(0.495)	1.065	(0.434)	-1.006	(0.592)	0.843
	PSD-E	-0.422	(1.098)	1.073	(0.488)	-0.976	(0.588)	1.903
500	OLS	-0.221	(0.236)	1.147	(0.192)	-1.12	(0.28)	0.181
	GLS	-0.262	(0.210)	1.153	(0.180)	-1.116	(0.261)	0.157
	PSD-I	-0.201	(0.242)	1.192	(0.205)	-1.171	(0.284)	0.182
	PSD-E	-0.232	(0.214)	1.172	(0.182)	-1.154	(0.265)	0.153
1000	OLS	-0.216	(0.168)	1.166	(0.135)	-1.155	(0.196)	0.088
	GLS	-0.233	(0.144)	1.171	(0.123)	-1.157	(0.180)	0.072
	PSD-I	-0.205	(0.171)	1.189	(0.142)	-1.182	(0.201)	0.090
	PSD-E	-0.220	(0.150)	1.177	(0.126)	-1.173	(0.187)	0.075
5000	OLS	-0.205	(0.076)	1.192	(0.058)	-1.189	(0.091)	0.018
	GLS	-0.203	(0.037)	1.196	(0.039)	-1.195	(0.05)	0.005
	PSD-I	-0.202	(0.076)	1.197	(0.061)	-1.196	(0.092)	0.018
	PSD-E	-0.200	(0.043)	1.197	(0.040)	-1.201	(0.058)	0.007

Notes: OLS and GLS are our closed-form estimators that are inefficient and efficient respectively. PSD-I and PSD-E are asymptotic least squares estimators of Pesendorfer and Schmidt-Dengler (2008) that are inefficient (identity weighted) and efficient respectively.

Table 3.3: Monte Carlo results (Equilibrium 3)

T	Estimator	F		θ_{10}		θ_{20}		MSE
100	OLS	-0.304	(0.475)	0.997	(0.398)	-0.895	(0.558)	0.840
	GLS	-0.436	(0.356)	1.015	(0.352)	-0.88	(0.446)	0.641
	PSD-I	-0.241	(0.514)	1.102	(0.471)	-1.023	(0.624)	0.917
	PSD-E	-0.397	(0.445)	1.081	(0.381)	-0.975	(0.526)	0.722
500	OLS	-0.225	(0.244)	1.149	(0.187)	-1.118	(0.282)	0.184
	GLS	-0.26 0	(0.229)	1.159	(0.185)	-1.122	(0.278)	0.175
	PSD-I	-0.201	(0.258)	1.200	(0.222)	-1.176	(0.304)	0.208
	PSD-E	-0.23	(0.239)	1.177	(0.189)	-1.157	(0.287)	0.178
1000	OLS	-0.214	(0.177)	1.169	(0.134)	-1.158	(0.204)	0.093
	GLS	-0.227	(0.170)	1.179	(0.136)	-1.166	(0.206)	0.092
	PSD-I	-0.202	(0.180)	1.193	(0.147)	-1.187	(0.211)	0.099
	PSD-E	-0.207	(0.186)	1.191	(0.148)	-1.188	(0.220)	0.105
5000	OLS	-0.203	(0.082)	1.194	(0.062)	-1.19	(0.093)	0.019
	GLS	-0.205	(0.076)	1.197	(0.060)	-1.192	(0.090)	0.017
	PSD-I	-0.201	(0.083)	1.200	(0.066)	-1.196	(0.095)	0.020
	PSD-E	-0.201	(0.078)	1.199	(0.061)	-1.197	(0.094)	0.018

Notes: OLS and GLS are our closed-form estimators that are inefficient and efficient respectively. PSD-I and PSD-E are asymptotic least squares estimators of Pesendorfer and Schmidt-Dengler (2008) that are inefficient (identity weighted) and efficient respectively.

The results are as expected from our theory. At smaller sample sizes the estimators are genuinely different regardless of the choice of weight matrices. Since the model is fully parametric both efficient estimators generally perform better than the inefficient ones even at $T = 100$ across all equilibria. With larger sample sizes the inefficient and efficient estimators seem to have similar properties for both methods. Although, in theory, the inefficient estimators need not be asymptotically equivalent as both are weighed by the same identity matrix (see the proof of Proposition E in the Appendix).

Table 3.4: Computation time

M	1	10	20	30	100	200
OLS	0.0021	0.0125	0.0245	0.0366	0.1241	0.2654
	(0.0010)	(0.0000)	(0.0000)	(0.0001)	(0.0004)	(0.0004)
GLS	0.0180	0.1542	0.3091	0.4658	1.8504	5.6084
	(0.0038)	(0.0001)	(0.0013)	(0.0002)	(0.0023)	(0.0069)
PSD-I	0.2084	4.9957	28.6415	73.3173	1171.5137	5657.6393
	(0.0089)	(0.0351)	(0.1805)	(0.0846)	(1.9478)	(0.9183)
PSD-E	0.3564	10.4140	52.0471	109.5519	1607.2349	7621.5963
	(0.0079)	(0.0359)	(0.1824)	(0.1049)	(2.6654)	(1.2093)

Notes: OLS and GLS are our closed-form estimators that are inefficient and efficient respectively. PSD-I and PSD-E are asymptotic least squares estimators of Pesendorfer and Schmidt-Dengler (2008) that are inefficient (identity weighted) and efficient respectively.

We now abstract away from the statistical properties and consider the numerical aspects. To illustrate the potential for computational advantages of our estimator, we introduce an additive market fixed effect to the per period payoff in the game described above. We use the number of markets, denoted by M, to control the complexity of the game.[7] For each M, we solve the model once and simulated five times using the symmetric equilibrium. We report in Table 4, the average central processing unit (CPU) times in seconds to compute our estimators and ALSE_{PSD} that minimize their respective limiting objective functions (no sampling error, using true choice and transition probabilities); standard errors are in parentheses.[8] Our estimators are substantially faster to compute, and the distinction grows

[7] There are other ways to vary the complexity of the game, e.g. by changing the number of potential actions and states. However, the difficulty to solve and estimate such games increases significantly as the games become more complexed. Our design only requires us to solve a simple game multiple times.

[8] The simulation was performed using MATLAB (R2012a, 64 bit version), on a standard PC running on an

exponentially with more parameters in the model. We also expect the computation time for ALSE_{PSD} to grow at a faster rate with larger action and/or state spaces for any fixed M. Another, perhaps even more, important numerical property of our closed-form estimators is they are always global minimizers. In contrast, a numerical solution to a general nonlinear optimization problem can be sensitive to the search algorithm, initial values, as well as the nature of the objective function.[9]

3.5 Conclusions and Possible Extensions

We have shown there can be some non-trivial computational gains in defining estimators that optimize objective functions constructed in terms of expected payoffs instead of choice probabilities for the estimation of structural dynamic discrete choice problems. The most transparent advantages of our approach follow from an opportunity to utilize familiar linear regression techniques, which arise when the period payoff functions are modeled to have fully or partially linear-in-parameter structure. Since the class of estimators we propose is asymptotically equivalent to the one developed by Pesendorfer and Schmidt-Dengler (2008), which includes several well-known estimators in the literature, there appears to be no costs at the theoretical level. Our estimators also perform well in Monte Carlo exercises in terms of speed and statistical properties.

The computation advantages we describe in this paper accumulates beyond the procedure to obtain a point estimate. For instance, resampling methods that are often used in practice to obtain standard errors (or perhaps to improve finite sample properties) clearly would benefit. The type of objective functions we propose also naturally complements other research in the literature that aims to improve the performance and/or scope of two-step methodologies. Two traditional criticisms of two-step estimators are large finite sample bias (from the first stage nonparametric estimation of choice probabilities), and the inability to accommodate unobserved heterogeneity and state variables that are persistent over time. For the former, Aguirregabiria and Mira (2002,2007) propose an iteration scheme that can improve the finite sample properties by imposing some structure for the first stage estima-

Intel Core (TM) 2 Duo 3.16 GHz processor with 4 GB RAM.

[9] It is easy to construct a game where the (limiting) objective function defined using pseudo-probabilities has multiple local minima such that some popular built-in optimization package, such as fminunc in MATLAB, produces different minimizers that depend on the initial search value.

tors; see Kasahara and Shimotsu (2008,2012) for further discussions and some theoretical justifications. At each iteration, the structural estimator can update the choice probabilities implied by the pseudo-model that are then used to define a new pseudo-likelihood function. To incorporate our estimator, alternatively one can use the updated probabilities to construct an objective function that defines the distance between the (updated) observed and implied expected payoffs. For the latter, the recent nonparametric identification results of Kasahara and Shimotsu (2009) and Hu and Shum (2012) show any two-step approach can also be readily applied to estimate a more general dynamic model than the one considered in this paper.

Bibliography

[1] Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2005), "Econometric Tools for Analyzing Market Outcome," *Handbook of Econometrics*, vol. 6, eds. J. Heckman and E. Leamer. North-Holland.

[2] Aguirregabiria, V., and P. Mira (2002): "Swapping Nested Fixed Point Algorithm: a Class of Estimators for Discrete Markov Decision Models," *Econometrica*, **70**, 1519-1543.

[3] Aguirregabiria, V. and P. Mira (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, **75**, 1-53.

[4] Aguirregabiria, V., and P. Mira (2010): "Dynamic Discrete Choice Structural Models: A Survey," *Journal of Econometrics,* **156**, 38-67

[5] Aguirregabiria, V. and J. Suzuki (2013): "Identification and Counterfactuals in Dynamic Models of Market Entry and Exit," *Working Paper, University of Toronto.*

[6] Amemiya, T. (1985): *Advanced Econometrics*, Harvard University Press.

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

[8] Bajari, P., C.L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, **75**, 1331-1370.

[9] Bajari, P., V. Chernozhukov, H. Hong and D. Nekipelov (2009): "Identification and Efficient Estimation of a Dynamic Discrete Game," *Working paper*, University of Minnesota.

[10] Chamberlain, G. (1982): "Multivariate Regression Models for Panel Data," *Journal of Econometrics*, **18**, 5-46.

[11] Collard-Wexler, A. (2013): “Demand Fluctuations and Plant Turnover in the Ready-Mix Concrete Industry,” *forthcoming in Econometrica.*

[12] Davidson, R. and J.G. MacKinnon (1993): *Estimation and Inference in Econometrics*, Oxford University Press.

[13] Egesdal, M., Z. Lai and C. Su (2013): “Estimating Dynamic Discrete-Choice Games of Incomplete Information,” *Working Paper, University of Chicago Booth School of Business.*

[14] Frisch, R. and F.V. Waugh (1933): “Partial Time Regressions as Compared with Individual Trends,” *Econometrica*, **1**, 387-401.

[15] Gourieroux, C. and A. Monfort (1995): *Statistics and Econometric Models: Volume 1, General Concepts, Estimation, Prediction and Algorithms*, Themes in Modern Econometrics, Cambridge University Press.

[16] Hansen, L.P. (1982): “Large Sample Properties of Generalized Method of Moments Estimators,” *Econometrica*, **50**, 1029 -1054.

[17] Hotz, V., and R.A. Miller (1993): “Conditional Choice Probabilities and the Estimation of Dynamic Models,” *Review of Economic Studies*, **60**, 497-531.

[18] Hotz, V., R.A. Miller, S. Smith and J. Smith (1994): “A Simulation Estimator for Dynamic Models of Discrete Choice,” *Review of Economic Studies*, **61**, 265-289.

[19] Hu, Y. and M. Shum (2012): “Nonparametric Identification of Dynamic Models with Unobserved State Variables,” *Journal of Econometrics*, **171**, 32-44.

[20] Kasahara, H. and K. Shimotsu (2008): “Pseudo-Likelihood Estimation and Bootstrap Inference for Structural Discrete Markov Decision Models,” *Journal of Econometrics*, **146**, 92-106.

[21] Kasahara, H. and K. Shimotsu (2009), “Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices,” *Econometrica*, **77**, 135-175.

[22] Kasahara, H. and K. Shimotsu (2012), “Sequential Estimation of Structural Models with a Fixed Point Constraint,” *Econometrica*, **80**, 2303-2319

[23] Li, Q. and J.S. Racine (2006): *Nonparametric Econometrics*, Princeton University Press.

[24] Magnac, M. and D. Thesmar (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, **70**, 801-816.

[25] Matzkin, R.L. (1991): "Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models," *Econometrica*, **59**, 1315–1327.

[26] Newey, W.K. and D.L. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," *Handbook of Econometrics*, vol. 4, eds. R.F. Engle and D. McFadden. North-Holland.

[27] Norets, A. and S. Takahashi (2013): "On the Surjectivity of the Mapping Between Utilities and Choice Probabilities," *Quantitative Economics*, **4**, 149–155.

[28] Pakes, A., M. Ostrovsky, and S. Berry (2007): "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Example)," *RAND Journal of Economics*, **38**, 373-399.

[29] Pesendorfer, M., and P. Schmidt-Dengler (2008): "Asymptotic Least Squares Estimator for Dynamic Games," *Review of Economics Studies*, **75**, 901-928.

[30] Ryan, S. P. (2012): "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, **80**, 1019-1061.

[31] Rust, J. (1994): "Structural Estimation of Markov Decision Processes," *Handbook of Econometrics*, vol. 4, eds. R.F. Engle and D. McFadden. North-Holland.

[32] Rust, J. (1996): "Numerical Dynamic Programming in Economics," *Handbook of Computational Economics*, vol. 1, eds. H.M. Aumann, D.A. Kendrick and J. Rust. Elsevier.

[33] Srisuma, S. (2012): "Minimum Distance Estimators for Dynamic Games," *Quantitative Economics*, **4**, 549-583.

[34] Srisuma, S. and O.B. Linton (2012): "Semiparametric Estimation of Markov Decision Processes with Continuous State Space," *Journal of Econometrics*, **166**, 320-341.

[35] Srisuma, S. (2013): "Identification in Markov Decision Models," *Working Paper, University of Cambridge.*

[36] Tamer, E. (2003), "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, **70**, 147–165.

Appendix

Proof of Theorems

Proof of Theorem 1. Under A1 to A3, $\mathcal{S}(\theta;\mathcal{W})$ has a well-separated minimum at θ_0. Let $\psi(\theta)=\mathcal{Y}-\mathcal{X}(\theta)$ and $\widehat{\psi}(\theta)=\widehat{\mathcal{Y}}-\widehat{\mathcal{X}}(\theta)$. Under A4, it follows that $\sup_{\theta\in\Theta}\|\psi(\theta)\|<\infty$ and $\sup_{\theta\in\Theta}\left\|\widehat{\psi}(\theta)-\psi(\theta)\right\|=o_p(1)$. Then through some tedious algebra, of repeatedly adding nulls and using properties of the matrix norm:

$$\begin{aligned}\widehat{\mathcal{S}}\left(\theta;\widehat{\mathcal{W}}\right)-\mathcal{S}(\theta;\mathcal{W}) &= \widehat{\psi}(\theta)^\top\widehat{\mathcal{W}}\widehat{\psi}(\theta)-\psi(\theta)^\top\mathcal{W}\psi(\theta)\\ &= 2\psi(\theta)^\top\mathcal{W}\left(\widehat{\psi}(\theta)-\psi(\theta)\right)+o_p\left(\left\|\widehat{\psi}(\theta)-\psi(\theta)\right\|\right),\end{aligned}$$

where the smaller order terms are uniform over Θ under A2 - A3.

Therefore $\sup_{\theta\in\Theta}\left|\widehat{\mathcal{S}}\left(\theta;\widehat{\mathcal{W}}\right)-\mathcal{S}(\theta;\mathcal{W})\right|=o_p(1)$, and consistency follows from a standard argument (e.g. see Newey and McFadden (1994)).■

Proof of Theorem 2. Under our assumptions, $\widehat{\theta}(\widehat{\mathcal{W}})$ satisfies the first order condition from differentiating (3.2.10) with respect to θ with probability tending to 1, i.e.

$$0=\left(\left.\frac{\partial\widehat{\mathcal{X}}(\theta)}{\partial\theta^\top}\right|_{\theta=\widehat{\theta}(\widehat{\mathcal{W}})}\right)^\top\widehat{\mathcal{W}}\left(\widehat{\mathcal{Y}}-\widehat{\mathcal{X}}(\widehat{\theta}(\widehat{\mathcal{W}}))\right)$$

holds with probability tending to 1. Since $\mathcal{Y}-\mathcal{X}(\theta_0)=0$, by adding nulls, we have

$$\begin{aligned}\widehat{\mathcal{Y}}-\widehat{\mathcal{X}}(\widehat{\theta}) &= \widehat{\mathcal{U}}+E_1+E_2\\ &= \widehat{\mathcal{U}}-\nabla_{\mathcal{X}}\left(\widehat{\theta}(\widehat{\mathcal{W}})-\theta_0\right)+o_p\left(\left\|\widehat{\theta}(\widehat{\mathcal{W}})-\theta_0\right\|\right),\end{aligned}$$

where $E_1=-\left(\mathcal{X}(\widehat{\theta}(\widehat{\mathcal{W}}))-\mathcal{X}(\theta_0)\right)$ and $E_2=\widehat{\mathcal{X}}(\widehat{\theta}(\widehat{\mathcal{W}}))-\widehat{\mathcal{X}}(\theta_0)-\left(\mathcal{X}(\widehat{\theta}(\widehat{\mathcal{W}}))-\mathcal{X}(\theta_0)\right)$, and the second equality follows from A5 after applying mean value expansions to the terms in E_1 and E_2 around θ_0. By adding nulls and using properties of matrix norm, since $\widehat{\theta}(\widehat{\mathcal{W}})=\theta_0+o_p(1)$, we also have $\left\|\left(\left.\frac{\partial\widehat{\mathcal{X}}(\theta)}{\partial\theta^\top}\right|_{\theta=\widehat{\theta}(\widehat{\mathcal{W}})}\right)^\top\widehat{\mathcal{W}}-\nabla_{\mathcal{X}}^\top\mathcal{W}\right\|=o_p(1)$ under A2 and A5. Therefore

$\widehat{\theta}(\widehat{\mathcal{W}})$ also satisfies

$$0 = \nabla_{\mathcal{X}}^{\top}\mathcal{W}\left(\widehat{\mathcal{U}} - \nabla_{\mathcal{X}}\left(\widehat{\theta}(\widehat{\mathcal{W}}) - \theta_0\right)\right) + o_p\left(\frac{1}{\sqrt{N}} + \left\|\widehat{\theta}(\widehat{\mathcal{W}}) - \theta_0\right\|\right),$$

with probability tending to 1. Then it follows that

$$\sqrt{N}\left(\widehat{\theta}(\widehat{\mathcal{W}}) - \theta_0\right) = \left(\nabla_{\mathcal{X}}^{\top}\mathcal{W}\nabla_{\mathcal{X}}\right)^{-1}\nabla_{\mathcal{X}}^{\top}\mathcal{W}\widehat{\mathcal{U}} + o_p(1).$$

An application of Slutsky's theorem gives the result.■

Proof of Theorem 3. The proof for part (i) is standard (e.g. see Theorem 3.2 of Hansen (1982)). We claim the optimal weighting matrix converges in the limit to Σ^{-1}. Let $B = \mathcal{W}\nabla_{\mathcal{X}}\left(\nabla_{\mathcal{X}}{}^{\top}\mathcal{W}\nabla_{\mathcal{X}}\right)^{-1}$ and $C = \Sigma^{-1}\nabla_{\mathcal{X}}\left(\nabla_{\mathcal{X}}{}^{\top}\Sigma^{-1}\nabla_{\mathcal{X}}\right)^{-1}$, so we have $\Omega_{\mathcal{W}} = B^{\top}\Sigma B$ and $\Omega_{\Sigma^{-1}} = C^{\top}\Sigma C$. Using simple algebra, it can be shown that $B^{\top}\Sigma B - C^{\top}\Sigma C = (B-C)^{\top}\Sigma(B-C) \geq 0$. For part (ii), it follows from the proof of Theorem 2 that we did not use any specific information on $\widehat{\mathcal{W}}$ beyond the fact that it has a positive definite probability limit.■

Representation Lemma

Proof of Lemma R. First we introduce some additional notations that build on the terms defined in Section 3.2.2. Let $v_{i,\theta_i}^{a} = \left(v_{i,\theta_i}(a,x^1),\ldots,v_{i,\theta_i}(a,x^J)\right)$ for all a, and $\mathbf{v}_{i,\theta_i} = \left(v_{i,\theta_i}^{0},\ldots,v_{i,\theta_i}^{K}\right)^{\top}$, so that $\mathbf{v}_{i,\theta_i}$ is a $J(K+1)$ −vector.

Let $\pi_{i,\theta_i}^{a_1\ldots a_I} = \left(\pi_{i,\theta_i}(a_1,\ldots,a_I,x^1),\ldots,\pi_{i,\theta_i}(a_1,\ldots,a_I,x^J)\right)$ for all $a_1,\ldots,a_I$, and $\pi_{i,\theta_i} = \left(\pi_{i,\theta_i}^{0\ldots0},\ldots,\pi_{i,\theta_i}^{K\ldots K}\right)^{\top}$, so that π_{i,θ_i} is a $J(K+1)^I$ −vector. For any k let: $\mathbf{I}_d$ denote an identity matrix of size d; $\mathbf{H}_i$ denote a block-diagonal matrix $\mathbf{diag}\left(H_i^0,H_i^1,\ldots,H_i^K\right)$, where H_i^a denotes a $J\times J$ matrix such that $\left(H_i^a\right)_{jj'} = \Pr\left[x_{t+1} = x^{j'}|x_t = x^j, a_{it} = a\right]$; $\mathbf{M} = \left(\mathbf{I}_{(K+1)^I}\otimes M\right)$ where $M = (\mathbf{I}_J - L)^{-1}$ and L denotes a $J\times J$ matrix such that $(L)_{jj'} = \beta\Pr\left[x_{t+1} = x^{j'}|x_t = x^j\right]$; $\mathbf{R} = \begin{bmatrix} P^{0\ldots0} & \cdots & P^{K\ldots K} \\ \vdots & \cdots & \vdots \\ P^{0\ldots0} & \cdots & P^{K\ldots K}\end{bmatrix}$ be a $J(K+1)^I$ by $J(K+1)^I$ matrix, where $P^{a_1\ldots a_I} = diag(P(a_1,\ldots,a_I|x^1),\ldots,P(a_1,\ldots,a_I|x^J))$ with $P(a_1,\ldots,a_I|x) = \Pr[\alpha_{1,\theta_1}(s_{it}) = a_1,\ldots,\alpha_{I,\theta_I}(s_{It}) = a_I|x_t =$

$x] = \prod_{j=1}^{I} P_j(a_j|x)$, and let $\mathbf{R}_i = \begin{bmatrix} P_{i0}^{0\ldots0} & \cdots & P_{i0}^{K\ldots K} \\ \vdots & \cdots & \vdots \\ P_{iK}^{0\ldots0} & \cdots & P_{iK}^{K\ldots K} \end{bmatrix}$ be a $J(K+1)$ by $J(K+1)^I$ matrix, where $P_{ik}^{a_1\ldots a_I} = diag(P_{ik}(a_1,\ldots,a_I|x^1),\ldots,P_{ik}(a_1,\ldots,a_I|x^J))$ with $P_{ik}(a_1,\ldots,a_I|x) = \Pr[\alpha_{1,\theta_1}(s_{it}) = a_1,\ldots,\alpha_{i-1,\theta_{i-1}}(s_{i-1t}) = a_{i-1},\alpha_{i,\theta_i}(s_{it}) = k,\alpha_{i+1,\theta_{i+1}}(s_{i+1t}) = a_{i+1},\alpha_{I,\theta_I}(s_{It}) = a_I|x_t = x] = P_i(k|x)\prod_{j\neq i}^{I} P_j(a_j|x)$.

Define $\Delta v_{i,\theta_i}^a = \left(v_{i,\theta_i}(a,x^1) - v_{i,\theta_i}(0,x^1),\ldots,v_{i,\theta_i}(a,x^J) - v_{i,\theta_i}(0,x^J)\right)$ for all $a>0$, and $\Delta\mathbf{v}_\theta = \left(\Delta v_{i,\theta_i}^1,\ldots,\Delta v_{i,\theta_i}^K\right)^\top$. Let $\mathbf{D}$ denote the $JK^I \times J(K+1)^J$ matrix that performs the transformation $\mathbf{D}\mathbf{v}_\theta = \Delta\mathbf{v}_\theta$. Lastly, let $\underline{v}_i^a = \left(\underline{v}_i(a,x^1),\ldots,\underline{v}_i(a,x^J)\right)$ for all a, and define $\underline{\mathbf{v}}_i = \left(\underline{v}_i^0,\ldots,\underline{v}_i^K\right)^\top$, so that $\Delta\underline{\mathbf{v}}_i = \mathbf{D}\underline{\mathbf{v}}_i$ is a $J(K+1)-$vector. Then (3.2.6) immediately follows.■

Asymptotic Equivalence of ALSEs

Proof of Proposition E. In the proof of this proposition we shall assume standard regularity conditions hold throughout (i.e. we assume inverse of matrices exist, expected payoffs and functions are bounded and continuously differentiable etc.). As seen from the proof of Theorem 2, under standard regularity conditions $\widehat{\theta}(\widehat{\mathcal{W}})$ satisfies

$$\widehat{\theta}(\widehat{\mathcal{W}}) = \theta_0 + \left(\nabla_{\mathcal{X}}^\top \mathcal{W} \nabla_{\mathcal{X}}\right)^{-1} \nabla_{\mathcal{X}}^\top \mathcal{W} \widehat{\mathcal{U}} + o_p\left(\frac{1}{\sqrt{N}}\right). \tag{3.5.1}$$

Next we introduce ALSE$_{PSD}$. It shall be useful to bear in mind the illustrative discussion in Section 3.2.1. We first define some additional notations that build on the terms defined in Section 3.2.3. Let $\mathbf{P} = \left(\mathbf{P}_1^\top,\ldots,\mathbf{P}_I^\top\right)^\top$ and $\mathbf{P}_\theta = \left(\mathbf{P}_{1,\theta_1}^\top,\ldots,\mathbf{P}_{I,\theta_I}^\top\right)^\top$. Similarly, let $\Delta\mathbf{v} = \left(\Delta\mathbf{v}_1^\top,\ldots,\Delta\mathbf{v}_I^\top\right)^\top$ and $\Delta\mathbf{v}_\theta = \left(\Delta\mathbf{v}_{1,\theta_1}^\top,\ldots,\Delta\mathbf{v}_{I,\theta_I}^\top\right)^\top$. Then, by Hotz and Miller's inversion there exists an invertible and continuously differentiable map Γ such that $\mathbf{P} = \Gamma(\Delta\mathbf{v})$ and $\mathbf{P}_\theta = \Gamma(\Delta\mathbf{v}_\theta)$. In particular

$$\begin{aligned} \mathbf{P} &= \left(\Gamma_1(\Delta\mathbf{v}_1)^\top,\ldots,\Gamma_I(\Delta\mathbf{v}_I)^\top\right)^\top, \text{ and} \\ \mathbf{P}_\theta &= \left(\Gamma_1(\Delta\mathbf{v}_{1,\theta_1})^\top,\ldots,\Gamma_I(\Delta\mathbf{v}_{I,\theta_I})^\top\right)^\top, \end{aligned}$$

where Γ_i is the inverse of Φ_i, which is defined in the text. Therefore, in terms of $\mathcal{Y}$ and $\mathcal{X}(\theta)$,

$$\Delta\mathbf{v} - \Delta\mathbf{v}_\theta = \mathcal{Y} - \mathcal{X}(\theta).$$

Thus $\mathbf{P}$ and $\mathbf{P}_\theta$ are also deterministic functions of the preliminary estimators (that we denoted by γ_0). We denote the estimators of $\mathbf{P}$ and $\mathbf{P}_\theta$ by $\widetilde{\mathbf{P}}$ and $\widehat{\mathbf{P}}_\theta$ respectively, and these estimators are constructed based on the same $\widehat{\gamma}$ that define $\widehat{\mathcal{X}}$ and $\widehat{\mathcal{Y}}$. Note that, although $\mathbf{P} = \mathbf{P}_{\theta_0}$, $\widetilde{\mathbf{P}}$ and $\widehat{\mathbf{P}}_{\theta_0}$ are generally different. An ALSE$_{PSD}$, denoted by $\widehat{\theta}_{PSD}(\widehat{\mathcal{V}})$, is defined as the minimizer of

$$\min_{\theta\in\Theta}\left(\widetilde{\mathbf{P}} - \widehat{\mathbf{P}}_\theta\right)^\top \widehat{\mathcal{V}}\left(\widetilde{\mathbf{P}} - \widehat{\mathbf{P}}_\theta\right),$$

for some $\widehat{\mathcal{V}}$ that converges in probability to positive definite matrix $\mathcal{V}$ (cf. equation (21) on page 915 in Pesendorfer and Schmidt-Dengler (2008)). Under appropriate regularity conditions, it is straightforward to show, analogous to our Theorem 2, that

$$\sqrt{N}\left(\widehat{\theta}_{PSD}(\widehat{\mathcal{V}}) - \theta_0\right) \xrightarrow{d} N\left(0, \Psi_{\mathcal{V}}\right).$$

For a first order asymptotic equivalence, it suffices to only consider the local asymptotic properties of ALSE$_{PSD}$ around θ_0. Let $\nabla_{\mathbf{P}}$ denote $\frac{\partial \mathbf{P}_\theta}{\partial\theta^\top}\big|_{\theta=\theta_0}$. An ALSE$_{PSD}$ satisfies

$$0 = -\nabla_{\mathbf{P}}^\top\mathcal{V}\left(\widetilde{\mathbf{P}} - \mathbf{P} - \left(\widehat{\mathbf{P}}_{\widehat{\theta}_{PSD}(\mathcal{V})} - \mathbf{P}_{\theta_0}\right)\right) + o_p\left(\frac{1}{\sqrt{N}}\right).$$

As the problem is smooth, it can be shown generally that the condition above simplifies further to

$$0 = -\nabla_{\mathbf{P}}^\top\mathcal{V}\left(\widetilde{\mathbf{P}} - \mathbf{P} - \left(\widehat{\mathbf{P}}_{\theta_0} - \mathbf{P}_{\theta_0} + \mathbf{P}_{\widehat{\theta}_{PSD}(\mathcal{V})} - \mathbf{P}_{\theta_0}\right)\right) + o_p\left(\frac{1}{\sqrt{N}}\right).$$

So that we have

$$\widehat{\theta}_{PSD}(\mathcal{V}) = \theta_0 + \left(\nabla_{\mathbf{P}}^\top\mathcal{V}\nabla_{\mathbf{P}}\right)^{-1}\nabla_{\mathbf{P}}^\top\mathcal{V}\left(\widetilde{\mathbf{P}} - \mathbf{P} - (\widehat{\mathbf{P}}_{\theta_0} - \mathbf{P}_{\theta_0})\right) + o_p\left(\frac{1}{\sqrt{N}}\right).$$

By chain rule $\nabla_{\mathbf{P}}$ equals $\nabla_\Gamma\nabla_{\mathcal{X}}$, where ∇_Γ denotes the Jacobian of Γ evaluated at $\Delta\mathbf{v}$, and

$\frac{\partial \Delta \mathbf{v}_\theta}{\partial \theta^\top}\Big|_{\theta=\theta_0}$ equals $\nabla_{\mathcal{X}}$. Thus, we can write

$$\begin{aligned}\widehat{\theta}_{PSD}(\mathcal{V}) &= \theta_0 + \left(\nabla_{\mathcal{X}}^\top \nabla_\Gamma^\top \mathcal{V} \nabla_\Gamma \nabla_{\mathcal{X}}\right)^{-1} \nabla_{\mathcal{X}}^\top \nabla_\Gamma^\top \mathcal{V} \left(\widetilde{\mathbf{P}} - \mathbf{P} - (\widehat{\mathbf{P}}_{\theta_0} - \mathbf{P}_{\theta_0})\right) + o_p\left(\frac{1}{\sqrt{N}}\right) \\ &= \theta_0 + \left(\nabla_{\mathcal{X}}^\top \nabla_\Gamma^\top \mathcal{V} \nabla_\Gamma \nabla_{\mathcal{X}}\right)^{-1} \nabla_{\mathcal{X}}^\top \nabla_\Gamma^\top \mathcal{V} \nabla_\Gamma \widehat{\mathcal{U}} + o_p\left(\frac{1}{\sqrt{N}}\right),\end{aligned}$$

where the last equality follows from linearizing $\widetilde{\mathbf{P}} - \mathbf{P} - (\widehat{\mathbf{P}}_{\theta_0} - \mathbf{P}_{\theta_0})$ in terms of $\widehat{\mathcal{Y}} - \widehat{\mathcal{X}}(\theta_0)$. By defining $\mathcal{W}_{\mathcal{V}} = \nabla_\Gamma^\top \mathcal{V} \nabla_\Gamma$, we have

$$\widehat{\theta}_{PSD}(\mathcal{V}) = \theta_0 + \left(\nabla_{\mathcal{X}}^\top \mathcal{W}_{\mathcal{V}} \nabla_{\mathcal{X}}\right)^{-1} \nabla_{\mathcal{X}}^\top \mathcal{W}_{\mathcal{V}} \widehat{\mathcal{U}} + o_p\left(\frac{1}{\sqrt{N}}\right). \tag{3.5.2}$$

Therefore, by comparing (3.5.1) and (3.5.2), $\widehat{\theta}_{PSD}(\mathcal{V})$ has the same asymptotic distribution as $\widehat{\theta}(\mathcal{W}_{\mathcal{V}})$. In particular, let $\mathcal{V}^*$ denote the efficient weighting matrix for ALSE$_{PSD}$ so that $\Psi_{\mathcal{V}^*} \leq \Psi_{\mathcal{V}}$ for any positive definite matrix $\mathcal{V}$. Therefore the efficient ALSE$_{PSD}$, denoted by $\widehat{\theta}_{PSD}^*$, has the same asymptotic distribution as $\widehat{\theta}(\mathcal{W}_{\mathcal{V}^*})$ with $\mathcal{W}_{\mathcal{V}^*} = \nabla_\Gamma^\top \mathcal{V}^* \nabla_\Gamma$. Then it must hold, by Theorem 3(i), that $\Omega_{\Sigma^{-1}} \leq \Psi_{\mathcal{V}^*}$ since $\Omega_{\Sigma^{-1}}$ is the lower variance bound. To complete the proof, an identical argument can be made in the reverse direction. It is easy to show that any $\widehat{\theta}(\mathcal{W})$ that satisfies (3.5.1) also has the same asymptotic distribution as $\widehat{\theta}_{PSD}(\mathcal{V}_{\mathcal{W}})$, where $\mathcal{V}_{\mathcal{W}} = \nabla_{\Gamma^{-1}}^\top \mathcal{W} \nabla_{\Gamma^{-1}}$ (cf. $\mathcal{W}_{\mathcal{V}}$), and $\nabla_{\Gamma^{-1}}$ denotes the Jacobian of Γ^{-1} evaluated at $\mathbf{P}$ (that equals $(\nabla_\Gamma)^{-1}$ by the inverse function theorem). We omit further details to avoid repetition. Thus, it follows that $\Psi_{\mathcal{V}^*} \leq \Omega_{\Sigma^{-1}}$, hence we can also conclude that $\Psi_{\mathcal{V}^*} = \Omega_{\Sigma^{-1}}$.

In summary:

$$\begin{aligned}\sqrt{N}\left(\widehat{\theta}(\mathcal{W}) - \theta_0\right) &= \sqrt{N}\left(\widehat{\theta}_{PSD}(\mathcal{V}_{\mathcal{W}}) - \theta_0\right) + o_p\left(\frac{1}{\sqrt{N}}\right) \text{ with } \mathcal{V}_{\mathcal{W}} = \nabla_{\Gamma^{-1}}^\top \mathcal{W} \nabla_{\Gamma^{-1}}, \quad (3.5.3) \\ \sqrt{N}\left(\widehat{\theta}_{PSD}(\mathcal{V}) - \theta_0\right) &= \sqrt{N}\left(\widehat{\theta}(\mathcal{W}_{\mathcal{V}}) - \theta_0\right) + o_p\left(\frac{1}{\sqrt{N}}\right) \text{ with } \mathcal{W}_{\mathcal{V}} = \nabla_\Gamma^\top \mathcal{V} \nabla_\Gamma,\end{aligned}$$

and $(\mathcal{V}, \mathcal{W})$ can be replaced by any consistent estimators $(\widehat{\mathcal{V}}, \widehat{\mathcal{W}})$. Therefore our estimator and ALSE$_{PSD}$ can always be constructed to have the same asymptotic distribution and achieve the same lower variance bound.■

Chapter 4

Bank Privatization and the Supply of Banking Services: Evidence from a Dynamic Structural Model

Co-authored with Fabio A. Miessi Sanches (University of São Paulo)

Abstract

This paper examines the effects of bank privatization on the supply of banking services in small markets. A dynamic game between the major Brazilian public and private banks is estimated. We show that profits of private banks are positively affected by the number of public and negatively affected by the number of private branches operating in Brazilian small markets. The structural model is used to examine the effects of the privatization of public banks on the number of bank branches in small markets. The counterfactual shows that the number of branches operating in small markets drops significantly after the privatization.

4.1 Introduction

The discussion about the existence of public, state owned banks has been prominent in the banking literature since the 1960's - see Barth, Caprio and Levine (2001) and La Porta, López-de-Silanes and Shleifer (2002) and Levy Yeyati, Micco and Panizza (2007).

In favour of public banks the following has been argued: (i) Public banks finance unprofitable but socially valuable investment projects; (ii) they foster financial development and (iii) they provide financial access to populations living in areas that are unattractive for private institutions. Critics of public banks argue that (i) they are used as political instruments, providing employment, credit, subsidies or other benefits in return for political assistance, and (ii) they crowd-out more efficient, more competitive private banks, slowing down the development of the financial system.

This paper examines the effects of public banks privatization on the supply of banking services. A dynamic entry game between the major Brazilian public and private banks is estimated. A counterfactual experiment is used to analyze how the privatization of public banks affects the supply of banking services in small isolated markets.

Three main conclusions emerge. First, public banks generate positive profit spill-overs for private banks; second, private banks crowd-out private competitors. Our estimates show that the entry of a public bank in a given market increases the return of a private incumbent by 1.2 percent and the entry of a private bank reduces the return of a private incumbent by 0.05 to 1 percent. Third, the counterfactual in which public banks are sold to private players shows that the total number of active branches operating in the long-run in a typical small market drops from 3 to 0.5 on average. To guarantee that, after privatization, all small municipalities would have at least one active branch the government should give a subsidy of approximately 8% on the operational costs of private branches. We can infer that the present cost of this policy would be approximately US$175,000[1] per market.

These findings have important policy implications in developing countries. In these countries a large fraction of the population has no access to the banking market. Yet the access to financial services generates positive effects in terms of poverty reduction and economic growth in disadvantaged areas (Burgess and Pande (2005) and Pascali (2012)).

Our estimates do not allow us to disentangle the details of the spill-over channels. Broadly speaking, our findings are consistent with public banks (i) having a monopoly over a number of important Federal funds and (ii) being driven by social, as opposed to strategic or market reasons. The first element guarantees a large volume of credit for small markets - see Feler (2012). The second induces product differentiation between public and private banks: Public and private banks target different clients - see Coelho, Melo and Rezende (2012). In this

[1] Approximately R$350,000.%

case, the amount of cheap credit and public transfers poured in by the public banks in small isolated municipalities shifts the demand for banking services, making these markets more attractive for private players. This effect induces the entry of private players.

There is little prior empirical evidence of the effects of public banks on the performance of the banking market. The evidence is mixed. La Porta, López-de-Silanes and Shleifer (2002) study a cross section of countries and show that the presence of public banks in the market is associated with poorly developed financial markets. They conclude that the higher the public ownership in the banking sector, the lower is the average growth of the ratio of private credit to GDP. Similar findings were obtained in Barth, Caprio and Levine (2001) who find that greater state ownership of banks with more poorly developed banks, nonbanks, and securities markets.

Levy Yeyati, Micco and Panizza (2007) extend the dataset used in La Porta, López-de-Silanes and Shleifer (2002) by including more controls and a longer period of time. They find that no robust conclusion can be drawn. The findings depend strongly on the definition of financial development, the estimator and the sample definition. They conclude that there is *"(...) no significant correlation between state-ownership of banks and credit to the private sector"* [2]. Detragiache, Tessel and Gupta (2008) confirm the findings in Levy Yeyati, Micco and Panizza (2007).

Cole (2007) finds that the nationalization of public banks in India increased the amount of credit in the market. Feler (2012) analyses the privatization of state banks in Brazil. His findings are close to the findings in Cole (2007).

This paper builds a dynamic entry game in which the major Brazilian public and private banks are the players. The dynamic structure of the model is rationalized by the existence of substantial entry costs in the market. At each period these players have information about the state variables and decide simultaneously to be active or not active in a given market by maximizing an inter-temporal profit function. Entrants pay a entry cost. We assume that the profit function of the major private players is asymmetrically affected by public and private competitors.

We use data from 1002 isolated markets in Brazil during 1988-2010 to estimate the decision rules for public and private banks. By relying on micro data from a single market, we are able to reduce the market heterogeneity present in cross country regressions, which, as reported

[2] Levy Yeyati, Micco and Panizza (2007).

by Levy Yeyati, Micco and Panizza (2007), causes important bias in the conclusions obtained by the existing literature. We recover the primitives of the game that are consistent with the estimated decision rules. The model is solved for the entry probabilities. The market equilibrium is evaluated under the counterfactual scenario where public banks are privatized. We report consistent *ex ante* estimates of the effects of this policy on market outcomes. Our model is valuable to predict policy changes.

Methodologically our paper is related to the empirical industrial organization literature that studies the estimation of dynamic games - see Aguirregabiria and Nevo (2010), Bajari, Hong and Nekipelov (2010) and Pesendorfer (2010) for a rich discussion on the topic. Applications that are similar to ours are also found in Pesendorfer and Schmidt-Dengler (2003) for small businesses in Austria, Dunne, Klimek, Roberts and Xu (2013) for dentists and chiropractors in the US, Gowrisankaran, Lucarelli, Schmidt-Dengler and Town (2010) for hospitals in the US, Collard-Wexler (2013) for the concrete industry in the US, Ryan (2012) for the cement industry in the US and Kalouptsidi (2013) for the shipping industry. Other applications include Maican and Orth (2012), Minamihashi (2012), Lin (2011), Fan and Xiao (2012), Nishiwaki (2010), Arcidiacono, Bayer, Blevins, and Ellickson (2012), Jeziorski (2012), Snider (2009), Suzuki (2012), Sweeting (2011) and Beresteanu, Ellickson and Misra (2010).

To estimate the model we use the alternative Asymptotic Least Squares (ALS) estimator developed in Sanches, Silva and Srisuma (2013). Sanches, Silva and Srisuma (2013) show that there can be substantial computational gains when the ALS estimator developed in Pesendorfer and Schmidt-Dengler (2008) is specified in terms of expected payoffs instead of choice probabilities. They also show that under the assumption of linear-in-the-parameters payoffs, the proposed estimator have the familiar OLS expression[3]. Differently from other popular estimation procedures for dynamic games (e.g. Hotz and Miller (1993), Aguirregabiria and Mira (2007), Bajari, Benkard and Levin (2007) and Pesendorfer and Schmidt-Dengler (2008), among others) this approach allows us to avoid the use of numerical methods. In doing so the computational burden is greatly reduced.

This paper is organized as follows. The next section describes our dataset and the Brazilian banking market. Section 4.3 shows reduced form evidence of competition between pub-

[3]Besides, Monte Carlo simulations in Sanches, Silva and Srisuma (2013) show that the closed-form estimator is much faster and have better small sample properties than the Asymptotic Least Squares estimator developed in Pesendorfer and Schmidt-Dengler (2008).

lic/private players. Sections 4.4 and 4.5 describe the theoretical model, the empirical model and our main results. Section 6 discusses the fitting of the empirical model and our counterfactual analysis. The last section concludes the paper.

4.2 Data and Institutional Background

The data comes from the Brazilian Central Bank and from the Brazilian Ministry of Labour. The Brazilian Central Bank database has followed the activities of all Brazilian banks since 1900. These data contain the opening and closing dates[4] and the name of the chain that operates each branch for all branches opened since 1900 in all Brazilian municipalities. A measure of market size is constructed by using data from the Brazilian Ministry of Labour containing the total payroll in the formal sector [5] for all Brazilian cities since 1985. The payroll data is deflated using the official inflation index, IPCA-IBGE. All the values are in R$ of 2011. The information about banking and economic activity is annual.

Following Bresnahan and Reiss (1991) our analysis examines small isolated markets. We select municipalities[6] that are at least 20 km away from the nearest municipality. State capitals and metropolitan areas are excluded. We also excluded municipalities that had more than 10 bank branches since 1900. This selection leaves us with 1002 isolated small markets, corresponding, roughly, to 20% of all Brazilian municipalites. Isolated markets enable us to obtain a clear measure of the potential demand for each branch.

The market size data starts in 1985. We exclude 1986 and 1987 from our sample because a major macroeconomic shock caused by two heterodox stabilization plans [7] severely disorganized the Brazilian economy in those years. Our final sample consists of observations for 1002 isolated municipalities in the period 1988-2010. The vast majority of municipalities have either one or none branch per chain. Our initial focus of analysis is therefore on entry and exit patterns. [8]

■ **Sample statistics.** The next table illustrates the basic statistics of our sample.

[4]For the branches that were closed.
[5]Number and wage of employees in the formal sector of the economy.
[6]From now on we use municipality/market interchangeably.
[7]Cruzado Plan in 1986 and Bresser Plan in 1987.
[8]In the municipalities that had more than one branch operated by the same chain, which correspond to less than 4% of the total number of municipalities and around 0.2% of our sample, we aggregated the branch level information for each player that had more than one branch in the same market. The exclusion of these municipalities does not change our results. Therefore we kept this information in the dataset.

Table 4.1: Basic Sample Statistics 1988-2010

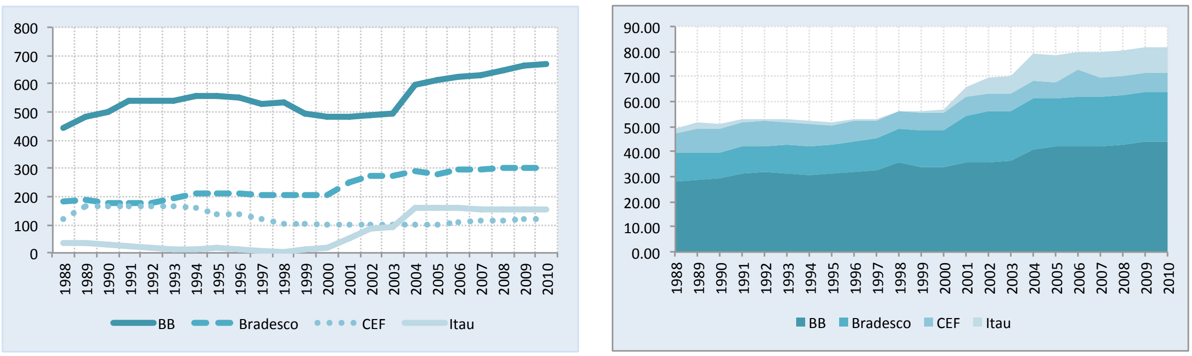
	Average	Std
Active Branches (**)	1560.2	142.4
Entry (**)	50.5	49.1
Exit (**)	59.5	55.1
Sample Market Size (*) (**)	10,404.85	
Municipalities	1002	
Sample Municipalities/Total Number of Municipalities	18%	
Municipalities × Periods	22041	

Note: * R$ millions of 2011. ** Yearly averages.

Our sample is composed by 1002 isolated markets. This corresponds to approximately 18% of the total number of municipalities in Brazil. The number of branches in this sample is 1560 per year on average. Entry is observed 50 times per year and exit 59 times. The yearly market size measured by the annual payroll of the formal workers of all the municipalities in the sample is of R$ 10.4 billions of 2011. This value is relatively small because by excluding state capitals and metropolitan regions, the richest cities in the country are left aside.

The Brazilian banking market is dominated by four big institutions: Two of them, Bank of Brazil, BB, and Caixa Economica Federal, CEF, are public and controlled by the federal government while the other two, Bradesco and Itau, are privately held. The next figures show the number of branches that are controled by these institutions and the market share, measured in terms of the number of active branches in the sample, of these four players.

Figure 4.1: Number of Branches (left) and Market Share (right) - "Big" Four



Note: Number of active branches per year (left) and fraction (right) of these branches over the total number of active branches in our sample.

These four players hold more than 80% of the active branches in our sample. The share of

Bradesco and Itau increased substantially over time. From 2000 to 2010 Bradesco's market share measured in terms of active branches increased from 13% to 20%. Itau's share increased from 1% to 10%. Part of the expansion is explainable by the acquisitions of privatized smaller public institutions. Bank of Brazil, BB, also experienced an increase in the number of branches. This expansion is mainly driven by the social policy of Lula's government (2003-2010), which tried to expand the presence of public banks in small markets.

Table 4.2: Average Monthly Payroll and Number of Public/Private Players

Number of Private	Number of Public	Mun Payroll	Observations
0	0	0.227	4592
0	1	0.478	6496
0	2	0.680	2511
1	0	0.515	1555
1	1	1.333	1696
1	2	1.677	1107
2	0	1.034	149
2	1	2.419	552
2	2	2.202	380

Note: Average market size is the monthly average payroll of the municipality and is measured in R$ millions of jan/2011 according to the number of players in the market. Sample period: 1988-2010. Each observation corresponds to a municipality in a given year. We showed in the table only the most frequent market structures. This corresponds to around 90% of the total number of observations.

Table 4.2 reports (i) the frequency distribution of each market configuration (number of observations corresponding to each market structure) and (ii) the average market size (monthly average payroll of the municipality in R$ millions of 2011) corresponding to each market structure. These numbers illustrate that:

1. Public players are located more frequently in small markets (as measured by the municipality average payroll) than private players; and,

2. Public players are frequently the only providers of financial services in these isolated markets (the frequency of public monopolies - 6496 observations - is the highest in the sample).

The Brazilian government has launched some programs that aim to "popularize" basic financial services in poorer areas, which includes the supply of basic services (current account, for example) and of credit lines to small farmers and firms. This feature may explain the fact that the frequency of markets where the public player is a monopolist (the only provider of financial services) is quite high in our sample. In addition, the empirical evidence indicates that public banks are much less productive than their private counterparts (Nakane and Weintraub (2005)). This means that the presence of public banks in smaller markets is not explained by cost advantages of public players.

■ **Institutional background.** The Brazilian banking market is large. In 2012 Itau was considered the 8th largest bank in the world in terms of market value (with a market value of US$88 billions); Bradesco was the 17th largest (market value of US$64 billions) and Bank of Brazil was the 31st largest (market value of US$42 billions)[9].

As pointed out in Coelho, Melo and Rezende (2012) there are important differences in the objectives of public and private banks. Private banks are essentially profit oriented. By legal mandate, public banks focus their operations on market segments that are not profitable for private banks. This suggests the existence of product differentiation in the market. In what follows we describe the "social" role of public banks in Brazil[10].

Bank of Brazil (BB) has expanded enormously its operations in smaller and poorer areas of the country based on central government policies aiming to "popularize" banking services among poor workers and small businesses. BB plays an important role as the provider of government funds to the Brazilian agriculture[11]. Also, to expand its capillarity in isolated areas, BB created a DSR (Regional Development Program). The DSR provides a set of tools for small entrepreneurs, including a business plan, technical support and credit [12].

Caixa Economica Federal (CEF) has a monopoly over a number of different government funds and services, such as the FGTS, Bolsa Família, PIS[13] and the Federal Lottery[14]. FGTS is a Brazilian fund created in 1966 to provide assistance to unemployed people [15] . These

[9] http://www.relbanks.com/worlds-top-banks/market-cap. Access: November 12, 2012.

[10] also present a detailed discussion on the role of Brazilian public banks.

[11] The total amount of agricultural credit provided by this player in 2010 reached more than US$ 26 billions. Moreover, BB is the main bank in the Pronaf, a program created to supply credit for small businesses (agriculture, fishing, turism, and handcraft) in rural areas at a very low interest rate. The total credit availble for the program increased from US$ 1 billion in 1999 to US$ 7 billions in 2010. All banks in Brazil are allowed to take part in the program, however, BB distributes around 65% of the total Pronaf credit.

[12] In 2007 the program supported 2800 business plans and distributed US$ 1.7 billions in credit.

[13] PIS is a tax to cover unemployment benefits. Their assets were around US$14 billions in December 2010.

[14] The Federal Lottery provided a gross revenue of US$5.2 billions in 2010. It is used to fund sports.

[15] The main source of funding is the monthly compulsory deposit that every private employer must do in the name of each employee. These values constitutes a fund and the worker have access to the money deposited in his/her name only in some special conditions: Unemployment, chronic disease, and for

resources are allocated in two main areas: Housing and sanitation. The government gives the investments guidelines in order to finance strategic areas with lack of credit. CEF is also responsible for the distribution of the benefits from Bolsa Família[16], a program that gives to poor families a monthly income. It was created to reduce the poverty in the most backward areas of the country.

Summarizing, the descriptive analysis suggests that the major Brazilian public banks have been used by the federal government to improve the financial access of isolated markets. The financial access includes new branches in smaller and poorer markets and the injection of cheap credit in these areas. These operations are not profitable for private banks. This suggests the existence of product differentiation in the market.

4.3 Reduced Form Analysis

We estimate a series of reduced form logit models to explain entry/exit movements of the biggest public and private players using the sample of isolated municipalities. We focus on the behavior of the 4 largest players: Bank of Brazil (BB), Caixa Economica Federal (CEF), Bradesco and Itau.

Two pooled logit models are estimated: One for the public players, BB and CEF; the other for private players, Itau and Bradesco. Our logit specification is:

$$\begin{gathered} P(a_{imt} = 1 | a_{imt-1}, n^{pub}_{mt-1}, n^{pri}_{mt-1}, \mathbf{x_{mt}}; \rho, \mu) = \\ \Lambda(\rho_0 + \rho_1 a_{imt-1} + \rho_2 n^{pub}_{mt-1} + \rho_3 n^{pri}_{mt-1} + \boldsymbol{\rho_4} \mathbf{x_{mt}} + \mu_t + \mu_m + \mu_{mt} + \mu_i) \end{gathered} \tag{4.3.1}$$

The dependent variable, a_{imt}, is the action of player i in municipality m, period t. It assumes 1 if player i was active in that municipality/period and zero otherwise. a_{imt-1} indicates the action of the same player in that municipality in the prior period. n^{pub}_{mt-1} is the

buying a house (if the worker does not own another house). CEF is responsible for the whole operation of the fund - from the tax collection to the payments for the benefited workers. Since the FGTS universe includes all formal workes (except public servants) the total size of the fund is considerably high - around US$90 billions in December 2008.

[16] In 2006 the program served around 11 million families or approximately 44 thousand individuals. The public expenditure with the program is around 0.5% of the Brazilian GDP and is growing steadily since its creation.

number of **public** competitors in the previous period in market m[17]; n^{pri}_{mt-1} is the number of **private** competitors in the previous period in market m[18]; $\mathbf{x_{mt}}$ is a vector of municipality characteristics; μ_t are time effects; μ_m are market specific effects; μ_{mt} captures market/time specific effects and μ_i are player specific effects. $\Lambda(\cdot)$ is the logistic distribution. The greek letters denote parameters to be estimated. The data include all muncipalities where player i was active for at least one period [19].

The estimation of dynamic binary response models with market fixed effects produces biased coefficients - see Carro (2007), Wooldridge (2010). To avoid this bias in our analysis, instead of including market dummies, 27 state dummies (one for each Brazilian state) were included in the model. States dummies are used to capture time invariant heterogeneity across municipalities of different states. Time effects are captured by year dummies. Time varying market effects are captured by the interaction of state dummies and a trend variable.

The vector $\mathbf{x_{mt}}$ includes municipality payroll, transfers of the Federal and State governments to the municipality, municipal government expenditure and agricutural production of the municipality. Municipality payroll is a measure of market size. The inclusion of transfers and municipal expenditure controls for the fact that entry of public banks can be correlated with an increase of Federal/State investment in the municipality, which can also affect entry of private banks. Agricultural production is included because a large fraction of the income in our isolated municipalities comes from agricultural activities. This variable is a different indicator of market size.

4.3.1 Private players

Table 4.3 reports the estimates of equation (4.3.1) for the private players, Bradesco and Itau. Only the marginal effects of n^{pub}_{mt-1} and n^{pri}_{mt-1} evaluated at the sample means are reported. The model fit is good, with Pseudo-R2 of 87%-91%. Strikingly, the number of public banks increases the entry probabilities of the private players by 10%-14%. The effects

[17] Mathematically, $n^{pub}_{mt-1} = \sum_{j \in i_{pub}, j \neq i} a_{jmt}$, where i_{pub} is the set of public players.

[18] Mathematically, $n^{pri}_{mt-1} = \sum_{j \in i_{pri}, j \neq i} a_{jmt}$, where i_{pri} is the set of private players.

[19] Our estimation approach is based on the potential markets for each player. The potential market is defined based on the super efficient estimator described in Pesendorfer and Schmidt-Dengler (2003). We defined that market m is a potential market for player i if $\underset{t}{max}\{a_{imt}, t = 1900, 1901, .., 2010\} = 1$, or, in other words, market m is a potential market for player i if she entered in that market at least for one period since 1900.

are very significant and robust across specifications. The inclusion of state dummies and the interaction between state dummies and the time trend increases this effect.

Interestingly, the number of private competitors reduces the entry probabilities of private players. This effect is around -6.4% in the specification with the full set of controls. It is statistically significant at 5%.

Table 4.3: Marginal Effects of n_{mt-1}^{pub} and n_{mt-1}^{pri} on the Entry Probabilities of Private Players, Bradesco and Itau

	(I)	(II)	(III)	(IV)
Nº Public	0.10544***	0.12839***	0.13829***	0.13237***
	[0.01]	[0.02]	[0.02]	[0.02]
Nº Private	-0.03532	-0.01930	-0.06206**	-0.06409**
	[0.02]	[0.03]	[0.03]	[0.03]
Player Dummy	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes
State Dummies	No	Yes	Yes	Yes
Trend*State Dummies	No	No	Yes	Yes
Transfers, Expenditure, Agric. Prod.	No	No	No	Yes
Observations	15,919	15,229	15,229	15,217
Pseudo R2	0.87	0.87	0.91	0.91

Note: (***) Significant at 1%; (**) significant at 5%; (*) significant at 10%. Marginal effects calculated at the sample means. Clustered standard errors in brackets. All the models have lagged activity, number of public and private competitors and municipality payroll. Transfers correspond to the total transfers of Federal and State governments to the municipality. Expenditure corresponds to municipal government expenditure. Agricultural Production is the total agricultural production of each municipality.

As a robustness check, the model was estimated in a subsample of municipalities that had (i) at least one public player and (ii) at least one and at most three public players in any time period. Tables 4.9 and 4.10 in the appendix report the results for each subsample.

The pattern of results remains unchanged when compared to the estimates in Table 4.3. This strategy is used to control for unobservable characteristics of markets with and without public players. In the subsample with at least one and at most three public players the market heterogeneity is reduced. Markets with roughly the same number of public players have similar observable characteristics.

4.3.2 Public players

Table 4.4 reports the estimates of equation (4.3.1) for the public players, Bank of Brasil and Caixa Economica Federal. The model fit is good, with Pseudo-R2 statistics around

85%. Strikingly, entry probabilities of public players are barely affected by the number of public and the number of private competitors in each market. Although significant in some specifications the marginal effects of the number of public and the number of private competitors are very small when contrasted with the estimates in Table 4.3. In all the specifications the marginal effects of n^{pub}_{mt-1} and n^{pri}_{mt-1} are small in magnitude, being below 1% and 0.5% respectively. In specification (IV) in Table 4.3 these effects were respectively 13% and -6.4%. These estimates imply that the marginal effects of n^{pub}_{mt-1} and n^{pri}_{mt-1} are around 13 times larger for private banks than for public banks. This pattern is robust to the inclusion of state dummies, the interaction of state dummies and the time trend variable, public transfers, municipal expenditure and agricultural production.

Table 4.4: Marginal Effects of n^{pub}_{mt-1} and n^{pri}_{mt-1} on the Entry Probabilities of Public Players, BB and CEF

	(I)	(II)	(III)	(IV)
Nº Public	0.00498***	0.00871***	0.00934***	0.00936***
	[0.00]	[0.00]	[0.00]	[0.00]
Nº Private	0.00016	0.00105	0.00455**	0.00455**
	[0.00]	[0.00]	[0.00]	[0.00]
Player Dummy	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes
State Dummies	No	Yes	Yes	Yes
Trend*State Dummies	No	No	Yes	Yes
Transfers, Expenditure, Agric. Prod.	No	No	No	Yes
Observations	20,357	20,357	20,357	20,350
Pseudo R2	0.83	0.84	0.87	0.87

Note: (***) Significant at 1%; (**) significant at 5%; (*) significant at 10%. Marginal effects calculated at the sample means. Clustered standard errors by municipality in brackets. All the models have lagged activity, number of public and private competitors and municipality payroll. Transfers correspond to the total transfers of Federal and State governments to the municipality. Expenditure corresponds to municipal government expenditure. Agricultural Production is the total agricultural production of each municipality.

4.4 Theoretical Model

This section sets up and solves a dynamic entry game between the major Brazilian banks. Motivated by the data, the game considers entry and exit decisions. In the data a chain has typically at most one branch in each municipality[20]. We focus on the behavior of two public

[20] As described above only in 4% of these municipalities one chain had more than 1 branch during the same period.

banks, Bank of Brazil and Caixa Economica Federal, and two private banks, Bradesco and Itau. In 2010, these players had more than 80% of the total number of active branches in our sample.

The model captures the features documented by the reduced forms. Dynamics can be rationalized by high entry costs[21]. Importantly, the model allows for different behavior of public and private players.

We estimate the primitives that rationalize the behavior of private banks using a dynamic oligopoly game. We do not structurally model the behavior of public banks. Entry decisions of public banks are assumed to do not depend on the number of public and the number of private competitors in the market. There are two explanations behind this assumption. First, the reduced form analysis suggests that entry probabilities of public banks are barely affected by the number of public and the number of private competitors in the market. Second, the literature recognizes that public banks are not necessarily profit maximizers. The behavior of public banks can depend on political and social reasons - see Levy Yeyati et al (2007), La Porta et al (2002) and Barth et al (2001).

At each period private players have information about the state variables and decide simultaneously to be active or not active in a given market by maximizing an inter-temporal profit function. Private players know that the entry of public banks do not depend on the actions of public and private competitors. Private entrants pay a entry cost. The profit function of the major private players is assumed to be asymmetrically affected by public and private competitors. This allows us to understand how public banks influence the performance of private players.

Closed related models were applied in Pesendorfer and Schmidt-Dengler (2003), Dunne, Klimek, Roberts and Xu (2013), Gorisankaran, Lucarelli, Schmidt-Dengler and Town (2010), Collard-Wexler (2013), Ryan (2012) and Kalouptsidi (2013), among others. Aguirregabiria and Nevo (2010), Bajari, Hong and Nekipelov (2010) and Pesendorfer (2010) present a rich discussion on the estimation of dynamic games.

[21] Market analysts point out that the returns of branches in small markets is quite low. Lower returns in these markets are explained by high fixed and operational costs and by reduced revenues - see Gonçalves and Sawaya (2007), Gouvea (2007) and Andrade (2007). This explains why the number of bank branches is small in the most backward areas of the country.

4.4.1 Assumptions

■ **Players.** There are two private players, Bradesco and Itau, which are indexed by $i_{pri} \in \{Bradesco, Itau\}$. There are two public players, Bank of Brazil, BB, and Caixa Economica Federal, CEF, which are indexed by $i_{pub} \in \{BB, CEF\}$[22].

■ **Time and markets.** Time is discrete, $t = 1, 2, ..., \infty$. There are $m \in \mathbf{M} = \{1, 2, 3, .., \overline{M}\}$ markets.

■ **Actions.** A player's action in market m, period t is denoted by $a_{im}^t \in \{0, 1\}$, where 0 means that a player is inactive; 1 means that a player is active. The $1 \times N$ vector $\mathbf{a_m^t}$ denotes the action profile in market m, period t. We sometimes use $\mathbf{a_{-im}^t}$ to denote the actions of all players but player i.

■ **State space.** The state space is discrete and finite. We use $\mathbf{s_m^t}$ to denote an element of the state space in market m. When necessary we use N_s to express the number of different possible states in market m.

■ **Transitions.** The vector $\mathbf{s_m^t}$ evolves according to the transition matrix $p_m(\mathbf{s_m^{t+1}}|\mathbf{s_m^t}, \mathbf{a_m^t})$, described by next period distribution of possible values for the vector $\mathbf{s_m^t}$ conditional on each possible current state and actions in municipality m. We sometimes use $\mathbf{p_m}$ to denote the vector of transitions, $p_m(\mathbf{s_m^{t+1}}|\mathbf{s_m^t}, \mathbf{a_m^t})$, for every possible future state $\mathbf{s_m^{t+1}}$ given any possible $(\mathbf{s_m^t}, \mathbf{a_m^t})$.

■ **Unobservables.** In each period players draw a profitability shock ε_{im}^t. The shock is privately observed while the distribution is publicly known.

■ **Payoffs of private players.** ***Private*** player's period payoff is:

$$\Pi(\mathbf{a_m^t}, \mathbf{s_m^t}; \mathbf{\Theta_{im}}) = \begin{cases} \pi_{im}(\mathbf{a_m^t}, \mathbf{x_m^t}) \\ +\mathbf{1}(a_{im}^t = 1)\varepsilon_{im}^t \\ +\mathbf{1}(a_{im}^t = 1)\mathbf{1}(a_{im}^{t-1} = 0)F_i. \end{cases} \tag{4.4.1}$$

Here $\pi_{im}(\mathbf{a_m^t}, \mathbf{x_m^t})$ denotes player $i's$ deterministic profits in market m, F_i are entry costs and ε_{im}^t is a profitability shock, which wil be specified below. $\mathbf{\Theta_{im}}$ denotes the parameters

[22] We also estimated a version of the model including a fringe of public/private players. The inclusion of these players does not change our results but increases substantially the state space of our model. This imposes computational difficulties to solve the model and to make conterfactual analysis. By this reason we do not include these players in the model.

in the model.

This specification captures the main aspects of our idea. An incumbent deciding to stay in the market receives period profits of $\pi_{im}(a_{im}^t = 1, \mathbf{a_{-im}^t}, \mathbf{x_m^t}) + \varepsilon_{im}^t$. An entrant receives the same profit as an incumbent but minus the sunk entry cost F_i. Any player that is outside and considers re-entering the market has to pay a fixed cost[23].

The term $\pi_{im}(\mathbf{a_m^t}, \mathbf{x_m^t})$ is a linear function of exogenous states and actions[24]

$$\pi_{im}(\mathbf{a_m^t}, \mathbf{x_m^t}) =$$
$$\left\{ \pi_{0i} + \pi_{1i}^{pub} \left(\sum_{j \in i_{pub}} a_{jm}^t \right) + \pi_{1i}^{pri} \left(\sum_{j \neq i, j \in i_{pri}} a_{jm}^t \right) + \pi_{2i} x_m^t \right\} \cdot \mathbf{1}(a_{im}^t = 1). \tag{4.4.2}$$

Here $\pi_{ji}^k \in \mathbb{R}^k$ are parameters and x_m^t is a demand shifter. This specification allows for different "competition" effects of public and private players.

The profitability shock ε_{im}^t is assumed to have three components:

$$\varepsilon_{im}^t = \mu_{im} + \eta_{it} + \xi_{im}^t,$$

where, μ_{im} is a term that varies only across markets and players but not over time, η_{it} is a time varying player specific term and $\xi_{im}^t \sim EV(0,1)$ is an idiosyncratic shock iid across individuals, time and markets. This is the only source of asymmetric information in the model. The first and the second elements of ε_{im}^t are known to the players and capture respectively (i) the correlation of the profitability shocks in the same market across time and (ii) correlation of the profitability shock across time in different markets. Both effects are empirically justified by the significance of state and year dummies in the reduced forms analyzed above.

The time varying shock is included to capture the fact that the decision structure of the chains can be centralized: First, the "general" conditions of the economy are observed; second the decision in which municipality(ies) to enter/exit is taken. The model captures the feature that a better (worse) macroeconomic landscape can increase (decrease) the probability of

[23]We assume that players leaving the market get a scrap value equal to zero. Aguirregabiria and Suzuki (2013) discuss identification problems of entry costs, scrap values and fixed costs in dynamic entry games.

[24]Similar structures were used in Pesendorfer and Schmidt-Dengler (2003, 2008), Ryan (2012) and Collard-Wexler (2013).

being active in all available markets.

We impose a structure on the time effect by further assuming $\eta_{it} = \eta_i \bar{x}_t$, where $\overline{x}_t = \sum_m x_{mt}$, is the total payroll of the municipalities in our sample in a given year. The process for the shock is:

$$\varepsilon_{im}^t = \mu_{im} + \eta_i \bar{x}_t + \xi_{im}^t$$

The parameters of interest are $\mathbf{\Theta_{im}} = \left\{ F_i, \pi_{0i}, \pi_{1i}^{pub}, \pi_{1i}^{pri}, \pi_{2i}, \mu_{im}, \eta_i \right\}$. We sometimes denote payoffs as $\Pi(\mathbf{a_m^t}, \mathbf{s_m^t}; \mathbf{\Theta_{im}}) = \tilde{\Pi}(\mathbf{a_m^t}, \mathbf{s_m^t})\mathbf{\Theta_{im}} + \mathbf{1}(a_{im}^t = 1)\xi_{im}^t$, where $\tilde{\Pi}(\mathbf{a_m^t}, \mathbf{s_m^t})$ is a $1 \times N_p$ vector and N_p is the number of parameters in market m.

We do not structurally model the behavior of public banks. Entry decisions of public banks are assumed to do not depend on the number of public and the number of private competitors in the market. We do not specify the payoff structure of public players. Our model accommodates all possible motivations for public banks' actions.

▪ **Sequence of period events.** The sequence of events of the game is the following:

1. States are observed by all the players.

2. Each player draws a private profitability shock ε_{im}^t.

3. Actions are simultaneously chosen. Players maximize their discounted sum of period payoffs given their beliefs on competitors' actions. The total payoff of a private player is given by the discounted sum of player's period payoffs. The discount rate is given by $\beta < 1$ and is the same for all players.

4. After actions are chosen the law of motion for $\mathbf{s_m^t}$ determines the distribution of states in the next period; the problem restarts.

Next the equilibrium for this game is characterized.

4.4.2 Equilibrium characterization

We restrict attention to pure *Markovian strategies*. This means that players' actions are fully determined by the current vector of state variables. Intuitively, whenever a player observes the same vector of states it will take the same actions and the history of the game until period t does not influence the player's decisions.

■ **Public players.** Public players' behavior are exogenously given. Entry probabilities of public players are known to the private players and do not depend on the actions of other public/private players.

■ **Private players.** Private player i's best response solves the following Bellman equation:

$$\underset{a_i^t=k\in\{0,1\}}{Max}\left\{\begin{array}{c}\sum_{\mathbf{a^t_{-im}}}\sigma_{im}(\mathbf{a^t_{-im}}|\mathbf{s^t_m})\Pi(a^t_{im}{=}k,\mathbf{a^t_{-im}},\mathbf{s^t_m};\mathbf{\Theta_{im}})+\\ \beta\mathbf{z_k}\left(\mathbf{s^{t+1}_m}|\mathbf{s^t_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)\mathbf{E}_\xi\mathbf{V_{im}}\left(\sigma_{\mathbf{im}},\mathbf{p_m}\right)\end{array}\right\}. \tag{4.4.3}$$

Here $\Pi(\cdot)$ is player's period payoff; the function $\sigma_{im}(\mathbf{a^t_{-im}}|\mathbf{s^t_m})$ accounts for i's beliefs on other players' actions given current states; $\sigma_{\mathbf{im}}$ is a vector that contains the beliefs for all possible combination of players actions given any possible state in market m; $\mathbf{z_k}\left(\mathbf{s^{t+1}_m}|\mathbf{s^t_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)$ is a $1\times N_s$ vector containing the transitions $\sigma_{im}(\mathbf{a^t_{-im}}|\mathbf{s^t_m})p_m(\mathbf{s^{t+1}_m}|a^t_{im}=k,\mathbf{a^t_{-im}},\mathbf{s^t_m})$ and $\mathbf{E}_\xi\mathbf{V_{im}}\left(\sigma_{\mathbf{im}},\mathbf{p_m}\right)$ is a $N_s\times 1$ vector with the expected continuation value for the player, $E_\xi V_i(\mathbf{s^{t+1}_m};\sigma_{\mathbf{im}},\mathbf{p_m},\mathbf{\Theta_{im}})$, for all $\mathbf{s^{t+1}_m}$.

The conditional value function, conditional on action $k\in\{0,1\}$ being played, is given by:

$$\begin{gathered}V^k_{im}\left(\mathbf{s^t_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)=\\ \sum_{\mathbf{a^t_{-im}}}\sigma_{im}(\mathbf{a^t_{-im}}|\mathbf{s^t_m})\tilde{\Pi}(a^t_{im}{=}k,\mathbf{a^t_{-im}},\mathbf{s^t_m})\mathbf{\Theta_{im}}\\ +\beta\mathbf{z_k}\left(\mathbf{s^{t+1}_m}|\mathbf{s^t_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)\mathbf{E}_\xi\mathbf{V_{im}}\left(\sigma_{\mathbf{im}},\mathbf{p_m}\right)+\mathbf{1}(k=1)\xi^t_{ik}.\end{gathered} \tag{4.4.4}$$

We also define $\tilde{V}^k_{im}\left(\mathbf{s^t_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)=V^k_{im}\left(\mathbf{s^t_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)-\mathbf{1}(k=1)\xi^t_{ik}$ as the conditional value function net of the iid profitability shock, $\mathbf{1}(k=1)\xi^t_{ik}$.

We define $\mathbf{E}_\xi\mathbf{V_{im}}\left(\sigma_{\mathbf{im}},\mathbf{p_m}\right)$ as the *ex-ante* value function, that is, $\mathbf{E}_\xi\mathbf{V_{im}}\left(\sigma_{\mathbf{im}},\mathbf{p_m}\right)=\mathbf{\Delta_{im}}\left(\tilde{\mathbf{\Pi}}_{\mathbf{im}}\mathbf{\Theta_{im}}+\tilde{\mathbf{E}}_{\xi\mathbf{im}}\right)$, where $\mathbf{\Delta_{im}}=\left[\mathbf{I_{N_s}}-\beta\mathbf{Z_{im}}\right]^{-1}$; $\tilde{\mathbf{\Pi}}_{\mathbf{im}}$ is a $N_s\times N_p$ vector stacking current payoff expected values, $\sum_{\mathbf{a^{t+1}_{im}}}\sigma_{im}(\mathbf{a^{t+1}_m}|\mathbf{s^{t+1}_m})\tilde{\Pi}(\mathbf{a^{t+1}_m},\mathbf{s^{t+1}_m})$, for every $\mathbf{s^{t+1}_m}$; $\tilde{\mathbf{E}}_{\xi\mathbf{im}}$ is a $N_s\times 1$ vector stacking $\tilde{E}_\xi\left(\mathbf{s^{t+1}_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)=\sum_{k=0}^{K}\sigma_{im}(a^{t+1}_{im}=k|\mathbf{s^{t+1}_m};\sigma_{\mathbf{im}},\mathbf{p_m})E\left[\xi^{t+1}_{im}|a^{t+1}_{im}=k,\mathbf{s^{t+1}_m}\right]$ for every $\mathbf{s^{t+1}_m}$; $\mathbf{I_{N_s}}$ is a $N_s\times N_s$ identity matrix; and $\mathbf{Z_{im}}$ is a $N_s\times N_s$ matrix stacking the $1\times N_s$ vector $\mathbf{z}\left(\mathbf{s^{t+2}_m}|\mathbf{s^{t+1}_m};\sigma_{\mathbf{im}},\mathbf{p_m}\right)$ containing the transitions $\sigma_{im}(\mathbf{a^{t+1}_m}|\mathbf{s^{t+1}_m})p_m(\mathbf{s^{t+2}_m}|\mathbf{a^{t+1}_m},\mathbf{s^{t+1}_m})$ for every $\mathbf{s^{t+1}_m}$.

The solution to problem (4.4.3) implies that player i's probability of playing action $k = 1$ when states are $\mathbf{s_m^t}$ satisfies the following equilibrium restrictions:

$$H_{im}(a_i^t = 1|\mathbf{s_m^t}; \sigma_{\mathbf{im}}, \mathbf{p_m}) = \\ 1 - exp\left\{-exp\left\{\ \tilde{V}_{im}^1\left(\mathbf{s_m^t}; \sigma_{\mathbf{im}}, \mathbf{p_m}\right) - \tilde{V}_{im}^0\left(\mathbf{s_m^t}; \sigma_{\mathbf{im}}, \mathbf{p_m}\right)\ \right\}\right\}. \tag{4.4.5}$$

This holds for all $\mathbf{s_m^t} \in \mathbf{S_m}$ and all $i \in i_{pri}$.

The solution to this problem is a vector of player i's optimal actions when the player faces each possible configuration for the state vector $\mathbf{s_m^t}$ and has consistent beliefs about other players actions in the same states of the world.

By stacking up best responses for every player and every state a system of $1 \times 2 \cdot N_s$ equations can be formed. This system is used to find the $1 \times 2 \cdot N_s$ vector of players' beliefs. A formal proof of the existence of this vector can be found in Pesendorfer and Schmidt-Dengler (2008). Equilibirum uniqueness, however, is not guaranteed. This is a common feature of entry games. The estimation procedure is designed to deal with the multiplicity of equilibria.

4.5 Econometric Model

This section describes identification and the estimation procedure. To estimate the model we use the estimator developed in Sanches, Silva an Srisuma (2013). Sanches, Silva and Srisuma (2013) show that there can be substantial computational gains when the ALS objective function developed in Pesendorfer and Schmidt-Dengler (2008) is specified in terms of expected payoffs instead of choice probabilities. They also show that under the assumption of linear-in-the-parameters payoffs, the proposed estimator have the familiar OLS expression. The estimator is easy to implement and reduces significantly the computational burden.

4.5.1 Identification

Following the CCP approach (Hotz and Miller (1993)) we firstly identify the vector of entry probabilities for public and private players and the transitions directly from the data. For the identification of entry probabilities we need to introduce two assumptions:

Assumption (i): *There are no unobserved common knowledge states.*

Assumption (ii): *The same equilibrium is played in all available markets.*

These identifying assumptions follow Ryan (2012). Pesendorfer (2010), Aguirregabiria and Nevo (2010) and Bajari, Hong and Nekipelov (2010) discuss the importance of the assumptions above.

Arcidiacono and Miller (2011) relax assumption (i). Our reduced form evidence suggests that our results are robust to the inclusion of market, time and player level unobservables. This mitigates our concern with unobservables.

Regarding assumption (ii), we could deal with the multiplicity problem by estimating the model for each market separately, as proposed by Pesendorfer and Schmidt-Dengler (2008). The main problem is that because we do not observe frequent entry and exit movements for a given market, we could not accurately identify the reduced form parameters. In our application it is necessary to pool the data of different markets. Our approach to pool data follows the earlier literature including Collard-Wexler (2013) and Ryan (2013).

Under these assumptions the identification of $\mathbf{\Theta_{im}}$ follows from Pesendorfer and Schmidt-Dengler (2008).

4.5.2 Estimator

We use the estimation principle developed in Sanches, Silva and Srisuma (2013) to recover players' payoffs. We start by representing the equilibrium restrictions (4.4.5) as a linear function of the payoff parameters. By inverting the function on the LHS of (4.4.5) and substituting $\tilde{V}^1_{im}\left(\mathbf{s^t_m}; \sigma_{\mathbf{im}}, \mathbf{p_m}\right)$ and $\tilde{V}^0_{im}\left(\mathbf{s^t_m}; \sigma_{\mathbf{im}}, \mathbf{p_m}\right)$ equation (4.4.5) can be written as:

$$y(\mathbf{s^t_m}; \sigma_{\mathbf{im}}, \mathbf{p_m}) - \mathbf{D}(\mathbf{s^t_m}; \sigma_{\mathbf{im}}, \mathbf{p_m})\mathbf{\Theta}^{'}_{\mathbf{im}} = 0, \tag{4.5.1}$$

where, $y(\mathbf{s^t_m}; \sigma_{\mathbf{im}}, \mathbf{p_m})$ is a real valued differentiable function that depends only on states, beliefs and state transitions and $\mathbf{D}(\mathbf{s^t_m}; \sigma_{\mathbf{im}}, \mathbf{p_m})$ is a real valued $1 \times N_p$ vector that depends only on states, beliefs and state transitions. These functions are defined in the appendix.

As in Pesendorfer and Schmidt-Dengler (2008) we assume that $\{(\hat{\sigma}_{\mathbf{im}}, \hat{\mathbf{p}}_{\mathbf{m}})\}_{m \in M}$ are consistent and asymptotically normally distributed estimators for the beliefs and state transitions in all the markets[25]. We define $\hat{y}_{imt} = y(\mathbf{s^t_m}; \hat{\sigma}_{\mathbf{im}}, \hat{\mathbf{p}}_{\mathbf{m}})$ and $\hat{\mathbf{D}}_{\mathbf{imt}} = \mathbf{D}(\mathbf{s^t_m}; \hat{\sigma}_{\mathbf{im}}, \hat{\mathbf{p}}_{\mathbf{m}})$ and sum

[25]As in Pesendorfer and Schmidt-Dengler (2008) we assume that consistency and asymptotic normality is

and subtract $\hat{y}_{imt} - \hat{\mathbf{D}}_{\mathbf{imt}}\mathbf{\Theta}^{'}_{\mathbf{im}}$ from (4.5.1) to write:

$$\hat{y}_{imt} = \hat{\mathbf{D}}_{\mathbf{imt}}\mathbf{\Theta}^{'}_{\mathbf{im}} + \hat{u}_{imt}, \tag{4.5.2}$$

where, $\hat{u}_{imt} = \left(\hat{y}_{ikmt} - \hat{\mathbf{D}}_{\mathbf{imt}}\mathbf{\Theta}^{'}_{\mathbf{im}}\right) - \left(y_{imt} - \mathbf{D}_{\mathbf{imt}}\mathbf{\Theta}^{'}_{\mathbf{im}}\right)$, with $y_{imt} = y(\mathbf{s}^{\mathbf{t}}_{\mathbf{im}}; \sigma_{\mathbf{im}}, \mathbf{p}_{\mathbf{m}})$ and $\mathbf{D}_{\mathbf{imt}} = \mathbf{D}(\mathbf{s}^{\mathbf{t}}_{\mathbf{m}}; \sigma_{\mathbf{im}}, \mathbf{p}_{\mathbf{m}})$.

By stacking equation (4.5.2) for all the markets, states and players OLS can be used to recover $\mathbf{\Theta}^{'}_{\mathbf{im}}$. Sanches, Silva and Srisuma (2013) show that the OLS estimator is consistent and asymptotically normally distributed when the number of observations used to compute $\{(\hat{\sigma}_{\mathbf{im}}, \hat{\mathbf{p}}_{\mathbf{m}})\}_{m \in M}$ tends to infinity.

4.5.3 CCPs and the state space

Following the CCP approach the empirical implementation of the model depends on (i) the estimation of beliefs and actions for each player, respectively, $H_{im}(a^t_i = 1|\mathbf{s}^{\mathbf{t}}_{\mathbf{m}}; \sigma_{im}(\cdot))$ and $\sigma_{im}(\mathbf{a}^{\mathbf{t}}_{-\mathbf{im}}|\mathbf{s}^{\mathbf{t}}_{\mathbf{m}})$ and (ii) the estimation of a transition process for the exogenous states, $p^s_m(\cdot)$. Next our estimation procedure for these elements is discussed.

Reduced form estimation of beliefs

We estimated equation (4.3.1) pooling the two private players, Bradesco and Itau. The data include the markets where Bradesco was active for at least one period and the markets where Itau was active for at least one period[26].

Instead of including year dummies we included $\overline{x}_t = \sum_m x_{mt}$, the total payroll of the municipalities in our sample in a given year, to control for the correlation in the decisions of private players in the same period of time. Instead of including state dummies we constructed 4 categories of markets. These market dummies keep the dimensionality of the state space low. Our market category definition follows the approach in Collard-Wexler (2013).

The market categories are defined according to the number of potential competitors in a given market[27]. More specificaly, if N_m is the number of potential competitors in municipality

obtained when the time series dimension of the observations used to compute $\{(\hat{\sigma}_{\mathbf{im}}, \hat{\mathbf{p}}_{\mathbf{m}})\}_{m \in M}$ goes to infinity.

[26] This follows the definition of potential markets defined in section 4.3.

[27] Our definition of potential competitor is based on the super efficient estimator in the section 4.3 - i.e. the number of potential competitors in municipality m is equal to the maximum number of players that were active for at least one period in municipality m since 1900.

m then $M_{1m} = 1$ if $N_m \leq 2$; $M_{2m} = 1$ if $2 < N_m \leq 4$; $M_{3m} = 1$ if $4 < N_m \leq 6$; and $M_{4m} = 1$ if $N_m \geq 7$. With this definition we can substitute $\mu_m = \sum_{k=1}^{4} \gamma_k M_{km}$. The same strategy was used in Collard-Wexler (2013)[28]. The vector $\mathbf{x_{mt}}$ includes only municipality payroll.

Entry decisions for public players were estimated using the same specification but excluding n_{mt-1}^{pub} and n_{mt-1}^{pri} from the set of covariates. We pooled the dara for the two public players, Bank of Brazil and Caixa Economica Federal. The data include all markets where either BB or CEF were active during at least one period.

We estimated the logits for the samples 1988-2010 and 1996-2010. The sample 1996-2010 excludes the hyperinflation period and allows us to focus on the more recent market trends. The logit model coefficient estimates for the public and private players are reported in the appendix.

State space and state transitions

Two distinct estimation strategies for the structural model parameters are explored: First, we exclude time and market effects as in Ryan (2013); second we use market dummies and the sample payroll to control for market and time effects. Both strategies are based on the empirical CCP estimates. Only the structural parameters for the private players, Bradesco and Itau, are estimated.

■ **Strategy 1: Model without time and market effects.** The state space for any private player, $i \in i_{pri}$, is composed by the following elements:

$$\mathbf{s_i^t} \in \left\{ a_i^{t-1}, \left\{ \mathbf{a_j^{t-1}} \right\}_{j \neq i}, x^t, \{\mathbf{I}(i = k)\}_{k \in i_{pri}} \right\}$$

Here $\{\mathbf{I}(i = k)\}_{k \in i_{pri}}$ is a set of private players dummies and x^t is the municipality payroll. The other elements are the actions of player i in period $t-1$, a_i^{t-1}, and the actions of player i's competitors in period $t-1$, $\left\{ \mathbf{a_j^{t-1}} \right\}_{j \neq i}$. The variable x^t is discretized in 10 deciles.

The law of motion for x^t is estimated by a simple auto-regressive ordered logit. This formulation for the law of motion of x^t ignores potential effects of banks, either public or private, on municipality income.

The state space of this model is composed by $2 \cdot 2^3 \cdot 10 \cdot 2 = 320$ elements.

■ **Strategy 2: Model with time and market effects.** The second model includes

[28] Collard-Wexler (2013) discusses potential endogeneity problems arising from the definition of the market dummies.

time and market effects. Market effects are captured by 4 market dummies. Time effects are captured by the sample payroll. Market dummies and the sample payroll variables were defined above - see section 4.5.3. The state space in this model is:

$$\mathbf{s^t_{im}} \in \left\{a^{t-1}_{im}, \left\{\mathbf{a^{t-1}_{jm}}\right\}_{j \neq i}, x^t_m, \bar{x}^t, \{\mathbf{I}(i=k)\}_{k \in i_{pri}}, \{\mathbf{I}(m=k)\}^4_{k=1}\right\}$$

Here $\{\mathbf{I}(m=k)\}^4_{k=1}$ is a set of market dummies for the 4 market types; $\bar{x}^t = \sum^4_{m=1} w_m x^t_m$ is the sample payroll, where, w_m is the number of markets of type m and x^t_m is the average payroll of type m markets in period t; a^{t-1}_{im} is player i's action in a market of type m in $t-1$, $\left\{\mathbf{a^{t-1}_{jm}}\right\}_{j \neq i}$ are the actions of player i's competitors in the same market in period $t-1$ and $\{\mathbf{I}(i=k)\}_{k \in i_{pri}}$ is a set of dummies for each private player.

The law of motion for x^t_m is calculated using an auto-regressive ordered logit structure. The variable x^t_m is discretized in four percentiles for each market type. A model for each market type was estimated. Finally $\bar{x}^t$ was calculated using $\bar{x}^t = \sum^4_{m=1} w_m x^t_m$ under the assumption that w_m is fixed over time.

The estimation of the model with market dummies is time consuming because the inclusion of the sample payroll, which depends on the realization of the payroll variable in every market, increases exponentially the dimension of the state space. The state space has $2 \cdot 2^3 \cdot 4^4 \cdot 2 \cdot 4 = 32768$ elements.

4.5.4 Results

We imposed an annual discount factor $\beta = 0.9$. To focus on the more recent trends of the market we used the CCPs and state transitions estimated with the 1996-2010 sample. The CCPs of the non strategic public players correspond to models I and II in the second block of Table 4.11 - those estimated using the sample 1996-2010. For the private players the CCPs are given by models I and II in the second block of Table 4.12. We used the OLS estimator to estimate the model.

Parameters are estimated in units of the scale factor in the EV distribution and do not have a level interpretation. Only relative magnitudes matter. Standard errors of the parameters were calculated by block bootstraping CCPs and state transitions 100 times. The structural model was estimated 100 times, one for each block bootstrap draw of beliefs and state transitions. The standard error across this set of parameters was calculated. A similar procedure was applied in Ryan (2012) and Collard-Wexler (2013).

Table 4.5 reports the structural parameters estimates. The first column corresponds to the model without market unobservables. The second column shows the model with market dummies and the sample payroll, estimated according to strategy 2.

Qualitatively, both specifications produce similar results. The main difference is that in model I, the market payroll coefficient is negative but small and not significant. All models predict that the entry of a new private competitor reduces the profits of the private incumbent. The entry of a new public player increases the profits of a private incumbent. The constant term, which measures operational costs, is negative and relatively larger in the second model. Entry costs are also negative and relatively larger in the second model. The contribution of the components of the shock in the second model is relatively important. The coefficient attached to the sample payroll is positive. This coefficient estimate means that increases in the sample income shifts to the right the distribution of the shock and increases entry rates. Market effects are positive.

Table 4.5: Structural Parameters for Private Players

	(I)	(II)
	Profit Components	
N Public	0.0605	0.0726
	[0.01]	[0.00]
N Private	-0.0487	-0.0256
	[0.01]	[0.00]
Market Payroll*	-0.0001	0.0019
	[0.00]	[0.00]
Constant	-0.3720	-0.5821
	[0.01]	[0.05]
	Shock Components	
Sample Payroll*		0.0004
		[0.00]
Market 1		0.2377
		[0.04]
Market 2		0.2104
		[0.03]
Market 3		0.1176
		[0.02]
	Entry/Player Costs	
Entry Costs	-4.9272	-5.7442
	[0.09]	[0.02]
Dummy Bradesco	-0.0270	-0.0245
	[0.01]	[0.02]
Observations	320	32768

Note: (*) Sample payroll measured in R$ billions of 2011; market payroll measured in R$ millions of 2011. Standard-errors in brackets. Standard errors obtained from 100 block bootstraps of beliefs and transitions. Parameters are measured in units of standard deviations of the iid profitability shock.

To facilitate the interpretation of these results, the next table reports the estimates as percentage of entry costs. All the coefficients are divided by the entry costs.

Table 4.6: Structural Parameters as Percentage of the Entry Costs

	(I)	(II)
	Profit Components	
N Public	1.228%	1.264%
N Private	-0.988%	-0.446%
Market Payroll*	-0.002%	0.033%
Constant	-7.550%	-10.134%
	Shock Components	
Sample Payroll*		0.007%
Market 1		4.138%
Market 2		3.663%
Market 3		2.047%
	Entry/Player Costs	
Entry Costs	100.000%	100.000%
Dummy Bradesco	-0.549%	-0.427%

Note: (*) Sample payroll measured in R$ billions of 2011; market payroll measured in R$ millions of 2011.

Again, the predictions of both models are quite close. Entry of a new public player increases profits of the private incumbent in around 1.3% of the entry costs. Entry of a new private player reduces profits of the private incumbent in around 0.45-0.9% of the entry costs.

The next table provides profit estimates for the private banks using the structural parameters. These parameters allow us to estimate a measure of return over entry costs. We also simulate the number of years necessary to recover entry costs.

Table 4.7: Average Period Profits and Return to Fixed Costs in Private Monopoly Markets

	(I)	(II)
Period Profits in Std Deviations	0.1936	0.2493
Period Profits as % of Entry Costs	3.930%	4.339%
Years to Recover Entry Costs	26.0	23.0

Note: Average profits of a private monopoly in a small market (market in the lower market payroll decile). Period profits as % of entry costs corresponds to the period profit in std deviations divided by the entry cost. To calculate the number of years to recover the entry costs we assumed a discount rate of 0.9, that the market payroll is increasing steadily at 3% per period and a monopoly structure every period.

The average period payoff of the private banks in monopoly markets is computed in the the first line of the table. The results show that the second model predicts larger profits. The second line shows that model I predicts returns to the entry investment of 4% in monopoly markets. Model II predicts that returns over entry costs are slightly larger, around 4.4%[29].

The third line shows the number of years necessary to recover the entry costs. We assumed $\beta = 0.9$ and that municipality and sample payroll are growing steadily at 3% per year. We accumulated the discounted payoffs and computed the number of years that are necessary to recover the estimated sunk entry costs. Model I predicts that in monopoly markets it takes on average 26 years for the private player to recover the sunk entry investment. Model II predicts that a private player needs on average 23 years to cover the sunk entry cost[30].

4.5.5 Discussion

Two remarkable facts arise from our analysis:

1. Public players complement private players;

2. Private players crowd-out other private players.

The first result shows that profit of public banks are positively affected by the number of public branches in the same market. Our estimates do not allow us to disentangle the details of the spill-over channels. Broadly speaking, our findings are consistent with public banks (i) having monopoly over a number of important Federal funds and (ii) being driven by social, as opposed to strategic or market reasons. The first element guarantees a large volume of credit for small markets - see Feler (2012). The second induces product differentiation between public and private banks: Public and private banks target different clients - see Coelho, Melo and Rezende (2012). In this case, the amount of cheap credit and public transfers poured in by the public banks into small isolated municipalities shifts the demand

[29] To calculate profits we fixed the sample payroll at its 2010 average value in smaller markets, that is in markets of type 1. For model II we assumed that the market dummy for markets of type 1 is equal to one. Thus the results are calculated for markets type 1. The sample payroll is used only to compute profits in model II and is equal to the sample payroll of 2010.

[30] This means that entry barriers are quite high. A recent expansion plan of Bank of Brazil illustrates this point. BB set down R$1 billion to construct 600 new branches in the Brazilian territory. This implies that on average each new branch costs R$1.66 million.

Notice also that the potential demand of a small market is quite small: The average yearly payroll of a market in our sample was R$9 million in 2010 and only a small fraction of the population demands banking services. In 2011 the Institute of Applied Economic Research (IPEA, 2011), an institute of the Brazilian federal government, estimated that around 40% of the Brazilian population has no access to any kind of banking services. This percentual can be even large in the markets represented in our sample.

for banking services, making these markets more attractive for private players. This effect induces the entry of private players.

The second result shows that private banks are competitive.

These findings contrast partially with the findings in Coelho, Mello and Rezende (2012). They extend the traditional Bresnahan and Reiss (1991) framework and estimate an entry model using a cross-section of small and medium Brazilian municipalities. This approach is related to ours, but in that paper the analysis is static and relies purely on cross-sectional variation[31]. They find that the presence of a private competitor reduces significantly profits of private incumbents. In contrast, the presence of a public competitor has a very small, but significant, negative effect on profits of private incumbents. They conclude that public banks are not competitive.

4.6 Model Fit and Counterfactual

This section uses the structural model to construct a policy experiment. We are interested in the following question: What happens with the supply of private financial services in small isolated markets when public banks are privatized?

First, we solve the model using the estimated parameters. The solution to the model is a vector of N_s entry probabilities that solves the system of implicit best responses given by equation (4.4.5).

For models with a large state space this exercise is not computationally feasible. The state space of model II has dimension $N_s = 32768$. Solving this model goes beyond current computational capabilities. The time to solve the model increases exponentially with the state space. From now on, we use only model I, that has a reduced state space ($N_s = 320$), to compute the counterfactual experiments.

4.6.1 Model fit

We solved the system (4.4.5) for private banks entry probabilities. This system is non linear. This means that its solution is not necessarily unique.

[31] The number of municipalities in Coelho, Mello and Rezende (2012) is substantially larger than ours. In their sample only state capitals and metropolitan areas are excluded. To construct our sample we select only municipalities that are at least 20 km away from the nearest municipality. State capitals and metropolitan areas are excluded.

To check how the multiplicity affects our conclusions, we proceed in the following way: First, we solve model I for the entry probabilities using the logit probabilities as the initial guess; second, we perturbed the logit probabilities; third we computed again the solution for the model using the "perturbed" vector of logit probabilities as the initial guess; fourth, we compared the "perturbed" solution with the original solution[32]. In doing so we find that the solutions were identical for any initial guess.

We compare the solution obtained from the structural model with the logit probabilities for all available states. The next table illustrates some statistics of our predictions.

Table 4.8: Fitted vs Sample (Logit) Probabilities

	(I)
Correlation Fitted and Logit Probabilities	99.91%
Average Sum of Squared Errors	0.05%
Average Sum of Errors	0.82%

Note: Correlation between the probabilities obtained from the solution of model I and the logit model (model I, Table 4.12 for 1996-2010) for each state (320 states). Average sum of squared errors gives the sum of the squared difference between the logit and the model probabilities for each state averaged across states. The average sum of errors gives the sum of the differences between the logit and the model probabilities averaged across the 320 states.

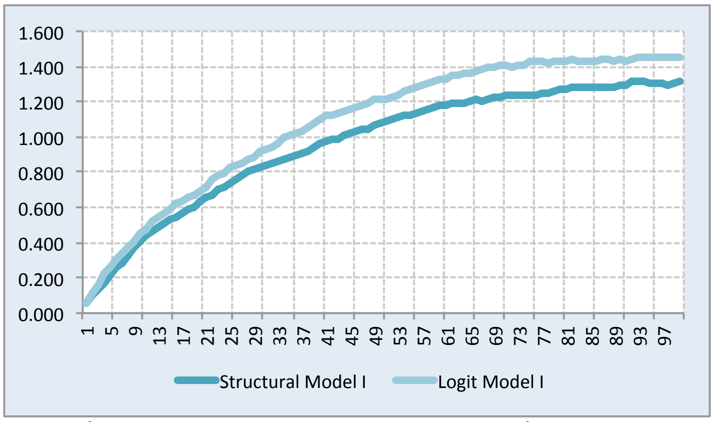
The first line reports the correlation between the logit probabilities and the solution of the structural model for all states. The second line reports the average squared difference between the logit and the structural probabilities. The third line reports the average sum of these differences. The fitting of the model is very good. The correlation between the logit and structural probabilities is high. The average error of the structural probabilities is below 1%.

We performed an additional exercise. We took the smallest market in terms of sample payroll and assumed that in the first period all the four banks are out of the market. We used firstly the probabilities predicted by the logit models and simulated 1000 paths 100 periods ahead of private banks actions and then we constructed an average path taking the mean across the 1000 paths. We did the same using the probabilities predicted by the structural models. The next figure compares the paths implied by the logit and by the structural

[32]Firstly we multiplied the original guesses (calculated from the CCPs showed above) by several factors between 0 and 1. We also started the model with a "fixed" guess, where the probabilities for all the states and for all the players are equal to 0.25, 0.5 and 0.75. We used the same procedure to compute the counterfactuals.

model.

Figure 4.2: Number of Private Banks 100 Periods Ahead - Model I



Note: Number of private banks in a small market starting from a state where all the competitors are out of the market. Paths 100 periods ahead simulated 1000 times using the structural and the logit probabilities for model I. The figure shows number of private branches averaged over 1000 simulations.

The figure shows that the path obtained from the structural model is close to the path obtained from the reduced form logit models in the first 30 periods. Subsequently the path of the structural model is below the path produced by the logit model. The structural model predicts that after 100 years this small market, without any public/private branch in operation in the first period, will have on average 1.32 private branches. The logit predicts that the same market will have 1.45 branches.

Next we use this model to construct counterfactuals.

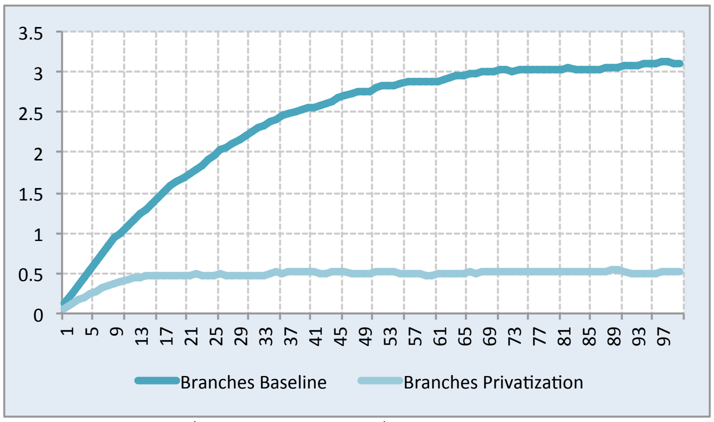
4.6.2 Counterfactual: Privatization of public banks

This section analyzes the effects of the privatization of both public banks on the total supply of financial services in small isolated markets. We assumed that each public bank is bought by different players: BB is bought by one player and CEF by the other. We assumed that the coefficient attached to the number of public competitors in the structural model is equal to the coefficient attached to the number of private competitors. The entry probabilities of public players, instead of being generated by an exogenous process, are calculated according to the system of best responses showed in equation (4.4.5).

We calculated the equilibrium probabilities for 4 players. Now, this calculation depends on the solution of a system of 640 equations and 640 unknown variables. To check how multiplicity affects our conclusions we used the procedure described in Section 4.6.1. In all experiments the resulting equilibrium did not change[33]. These probabilities are used to simulate 1000 paths 100 periods ahead. The next figure shows the path for the total number of branches, public plus private, after and before the privatization. We computed this path for a small market where the initial state is characterized by zero active players.

The exercise shows that in the long-run the total number of active branches in small municipalities drops from 3 to 0.5 on average. This means that with the privatization around 50% of the Brazilian small municipalities would not be attended by any bank branch. To assure that all these small municipalities would have at least one bank branch in the counterfactual world where public banks are bought by strategic players the government should give a subsidy of 8% over the operational costs of all active branches in the market. Using the fact that the structural model predicts that operational costs are around 7.55% of entry costs and an estimate of R$1 million for the entry costs we calculated that the present cost of this policy is around R$349,463.51 per municipality[34].

Figure 4.3: Counterfactual: Privatization of Public Players



Note: Number of branches (public plus private) in a small market starting from a state where all the competitors are out of the market (baseline and privatization counterfactual). Paths 100 periods ahead simulated 1000 times using the structural probabilities for model I. The figure shows number of private branches averaged over 1000 simulations. Branches privatization shows the total number of branches if public branches are privatized. Branches baseline is the total number of branches (public plus private) using the structural model I for private players and the non strategic behavior assumption for public players (calculated based on the logits in Table 4.11, model I, sample 1996-2010).

[33]Locally it is expected that the equilibrium is unique as the number of equilibria is generically finite. Thus equlibria are isolated points.

[34]Present values for a time horizon of 100 years and using a discount factor of 0.9 per year.

4.7 Conclusions

This paper explores microdata of 1002 isolated markets in Brazil to estimate a dynamic entry game for public and private banks. We recover players' payoffs. The model is solved for the equilibrium entry probabilities. The market equilibrium is evaluated under the counterfactual scenario where public banks are privatized.

Three main conclusions emerge. First, public banks generate positive profit spill-overs for private banks; second, private banks crowd-out private competitors. Our estimates show that the entry of a public bank in a given market increases the return of a private incumbent by 1.2 percent and the entry of a private bank reduces the return of a private incumbent by 0.05 to 1 percent. Third, the counterfactual in which public banks are sold to private players shows that the total number of active branches operating in the long-run in a typical small market drops from 3 to 0.5 on average.

Bibliography

[1] Aguirregabiria, V. and A. Nevo (2010): "Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games". *Advances in Economics and Econometrics*, Vol. III, Cambridge University Press.

[2] Aguirregabiria, V. and J. Suzuki (2013): "Identification and Counterfactuals in Dynamic Models of Market Entry and Exit," *Working Paper, University of Toronto.*

[3] Andrade, L. (2007): "Uma Nova Era para os Bancos da América Latina". *The Mckinsey Quarterly*, Edição Especial 2007: Criando uma Nova Agenda para a América Latina.

[4] Arcidiacono, P., P. Bayer, J. Blevins, and P. Ellickson (2012): "Estimation of Dynamic Discrete Choice Models in Continuous Time". *Working Paper, NBER.*

[5] Arcidiacono, P. and R. Miller (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity". *Econometrica*, **79(6)**, 1823-1867.

[6] Bajari, P., C.L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, **75**, 1331-1370.

[7] Bajari, P., H. Hong and D. Nekipelov (2010): "Game Theory and Econometrics: A Survey of Some Recent Research". *Advances in Economics and Econometrics*, Vol. III, Cambridge University Press.

[8] Bajari, P., H. Hong, J. Krainer and D. Nekipelov (2010): "Estimating Static Models of Strategic Interaction". *Journal of Business and Economic Statistics*, **28(4)**, 469-482.

[9] Barth, J. R. , G. Caprio and R. Levine (2001): "Bank Regulation and Supervision: What Works Best?" *Policy Research Working Paper, The World Bank.*

[10] Beresteanu, A., P. Ellickson and S. Misra (2010): “The Dynamics of Retail Oligopolies”. *Working Paper, University of Rochester, Simon School of Business.*

[11] Blundell, R. and S. Bond (1998): “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models”. *Journal of Econometrics*, **87(1)**, 115-143.

[12] Bresnahan, T. F. and P. C. Reiss (1991): “Entry and Competition in Concentrated Markets”. *Journal of Political Economy*, **99(5)**, 977-1009.

[13] Burgess, R. and R. Pande (2005): “Do Rural Banks Matter? Evidence from the Indian Social Banking Experiment”. *American Economic Review*, **95(3)**, 780-795.

[14] Carro, J. (2007): “Estimating Dynamic Panel Data Discrete Choice Models with Fixed Effects”. *Journal of Econometrics*, **140(2)**, 503-528.

[15] Coelho, C., J. Mello and L. Rezende (2012): “Are Public Banks pro-Competitive? Evidence from Concentrated Local Markets in Brazil”. *Forthcoming Journal of Money, Credit and Banking.*

[16] Cole, S. (2007): “Financial Development, Bank Ownership, and Growth. Or, Does Quantity Imply Quality?”. *Review of Economics and Statistics*, **91(1)**, 33-55.

[17] Collard-Wexler, A. (2013): “Demand Fluctuations and Plant Turnover in the Ready-Mix Concrete Industry,” *forthcoming in Econometrica.*

[18] Detragiache, E., T. Tressel and P. Gupta (2008): “Foreign Banks in Poor Countries: Theory and Evidence”. *Journal of Finance*, **63(5)**, 2123-2160.

[19] Dunne, T., S. Klimek, M. J. Roberts and Y. D. Xu (2013): “Entry, Exit, and the Determinants of Market Structure”. Forthcoming RAND Journal of Economics.

[20] Fan, Y. and M. Xiao (2012): “Competition and Subsidies in the Deregulated U.S. Local Telephone Industry”. *Working Paper, University of Michigan, Department of Economics.*

[21] Feler, L. (2012): “State Bank Privatization and Local Economic Activity”. *Working Paper, Johns Hopkins University, School of Advanced International Studies.*

[22] Gonçalves, L. and A. Sawaya (2007): "Financiando os Consumidores de Baixa Renda na América Latina". *The Mckinsey Quarterly*, Edição Especial 2007: Criando uma Nova Agenda para a América Latina.

[23] Gowrisankaran, G., C. Lucarelli, P. Schmidt-Dengler and R. Town (2010): "Government Policy and the Dynamics of Market Structure: Evidence from Critival Access Hospitals". *Eleventh CEPR Conference on Applied Industrial Organization, Toulouse School of Economics.*

[24] Gourieroux, C. and A. Monfort (1995): *Statistics and Econometric Models: Volume 1, General Concepts, Estimation, Prediction and Algorithms*, Themes in Modern Econometrics, Cambridge University Press.

[25] Gouvea, A. (2007): "Uma História de Sucesso no Setor Bancário da América Latina: Entrvista com Roberto Setúbal, CEO do Banco Itaú". *The Mckinsey Quarterly*, Edição Especial 2007: Criando uma Nova Agenda para a América Latina.

[26] Hotz, V., and R.A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, **60**, 497-531.

[27] Hotz, V., R.A. Miller, S. Smith and J. Smith (1994): "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, **61**, 265-289.

[28] Jeziorski, P. (2012): "Estimation of Cost Efficiencies from Mergers: An Application to US Radio". *Working Paper, UC Berkeley, Haas School of Business.*

[29] Kalouptsidi, M. (2013): "Time to Build and Fluctuations in Bulk Shipping". *Forthcoming American Economic Review.*

[30] La Porta, R., F. López-de-Silanes and A. Shleifer (2002): "Government Ownership of Banks". *Journal of Finance*, **57(1)**, 265-301.

[31] Levy Yeyati, E., A. Micco and U. Panizza (2007): "A Reappraisal of State-Owned Banks". *Economia*, **7(2)**, 209-247.

[32] Lin, H. (2011): "Quality Choice and Market Structure: A Dynamic Analysis of Nursing Home Oligopolies". *Working Paper, Indiana University, Kelley School of Business.*

[33] Maican, F. and M. Orth (2012): “Store Dynamics, Differentiation and Determinants of Market Structure”. *Working Paper, University of Gothenburg, Department of Economics.*

[34] Minamihashi, N. (2012): “Natural Monopoly and Distorted Competition: Evidence from Unbundling Fiber-Optic Networks”. *Working paper, Bank of Canada.*

[35] Nakane, M. I. and D. B. Weintraub (2005): “Bank Privatization and Productivity: Evidence for Brazil”. *Journal of Banking and Finance*, **29**, 2259-2289.

[36] Nishiwaki, M. (2010): “Horizontal Mergers and Divestment Dynamics in a Sunset Industry”. *Working Paper, National Graduate Institute for Policy Studies.*

[37] Pascali, L. (2012): “Banks and Development: Jewish Communities in the Italian Renaissance and Current Economic Performance”. *Barcelona GSE Working Paper.*

[38] Pesendorfer, M. (2010): “Estimation of (Dynamic) Games: A Discussion”. *Advances in Economics and Econometrics*, Vol. III, Cambridge University Press.

[39] Pesendorfer, M. and P. Schmidt-Dengler (2003): “Identification and Estimation of Dynamic Games”. *Working Paper, NBER.*

[40] Pesendorfer, M., and P. Schmidt-Dengler (2008): “Asymptotic Least Squares Estimator for Dynamic Games,” *Review of Economics Studies*, **75**, 901-928.

[41] Ryan, S. P. (2012): “The Costs of Environmental Regulation in a Concentrated Industry,” *Econometrica*, **80**, 1019-1061.

[42] Snider, C. (2009): “Predatory Incentives and Predation Policy: The American Airlines Case”. *Working Paper, UCLA, Department of Economics.*

[43] Sanches, F. A. M., D. Silva-Junior and S. Srisuma (2013): “An Alternative Asymptotic Least Squares Estimator for Dynamic Games”. *Working Paper, LSE, Department of Economics.*

[44] Sweeting, A. (2011): “Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry”. *Working Paper, Duke University, Department of Economics.*

[45] Suzuki, J. (2012): "Land Use Regulation as a Barrier to Entry: Evidence from the Texas Lodging Industry". *Forthcoming International Economic Review.*

[46] Wooldridge, J. M. (2010): "Econometric Analysis of Cross Section and Panel Data". *The MIT Press, Cambridge, MA.*

[47] World's Largest Banks 2012 (2012), viewed November 12 2012, http://www.relbanks.com/worlds-top-banks/market-cap.

Appendix 1: Proofs

■ $y(\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m})$ **and** $\mathbf{D}(\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m})$ **functions.**

The functions $y(\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m})$ and $\mathbf{D}(\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m})$ in equation (4.5.1) are defined as:

$$\begin{gathered}
y(\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m}) = \\
ln\left\{ln\left[\frac{1}{1-H_{im}(a_i^t=1|\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m})}\right]\right\} - \\
\beta\left[\mathbf{z_1}\left(\mathbf{s_m^{t+1}}|\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m}\right) - \mathbf{z_0}\left(\mathbf{s_m^{t+1}}|\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m}\right)\right]\mathbf{\Delta_{im}}\tilde{\mathbf{E}}_{\xi\mathbf{im}},
\end{gathered}$$

which is a real valued differentiable function that depends only on states, beliefs and transitions and,

$$\begin{gathered}
\mathbf{D}(\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m}) = \\
\sum_{\mathbf{a^t_{-im}}}\sigma_{im}(\mathbf{a^t_{-im}}|\mathbf{s_m^t})\tilde{\Pi}(a_{im}^t=1,\mathbf{a^t_{-im}},\mathbf{s_m^t}) + \\
\beta\left[\mathbf{z_1}\left(\mathbf{s_m^{t+1}}|\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m}\right) - \mathbf{z_0}\left(\mathbf{s_m^{t+1}}|\mathbf{s_m^t};\sigma_{\mathbf{im}},\mathbf{p_m}\right)\right]\mathbf{\Delta_{im}}\tilde{\mathbf{\Pi}}_{\mathbf{im}},
\end{gathered}$$

which is a real valued $1\times N_p$ differentiable vector that depends only on states, beliefs and transitions.

Appendix 2: Reduced Form Estimates

Table 4.9: Marginal Effects of n^{pub}_{mt-1} and n^{pri}_{mt-1} on the Entry Probabilities of Private Players (Bradesco and Itau) - Subsample $n^{pub} \geq 1$

	(I)	(II)	(III)	(IV)
Nº Public	0.11467***	0.12303***	0.15040***	0.14810***
	[0.02]	[0.02]	[0.03]	[0.03]
Nº Private	-0.04343*	-0.02822	-0.07914**	-0.07935**
	[0.02]	[0.03]	[0.03]	[0.03]
Player Dummy	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes
State Dummies	No	Yes	Yes	Yes
Trend*State Dummies	No	No	Yes	Yes
Transfers, Expenditure, Agric. Prod.	No	No	No	Yes
Observations	9,348	9,164	9,164	9,162
Pseudo R2	0.87	0.88	0.92	0.92

Note: (***) Significant at 1%; (**) significant at 5%; (*) significant at 10%. Marginal effects calculated at the sample means. Clustered standard errors at the municipality level in brackets. All the models have lagged activity, number of public and private competitors and municipality payroll. Subsample $n^{pub} \geq 1$ includes all municipalities that had at least one public player in every period.

Table 4.10: Marginal Effects of n^{pub}_{mt-1} and n^{pri}_{mt-1} on the Entry Probabilities of Private Players (Bradesco and Itau) - Subsample $1 \leq n^{pub} \leq 3$

	(I)	(II)	(III)	(IV)
Nº Public	0.19269***	0.18806***	0.20575***	0.21391***
	[0.03]	[0.04]	[0.04]	[0.04]
Nº Private	-0.05677*	-0.06105*	-0.12490***	-0.11986***
	[0.03]	[0.04]	[0.05]	[0.05]
Player Dummy	Yes	Yes	Yes	Yes
Time Dummies	Yes	Yes	Yes	Yes
State Dummies	No	Yes	Yes	Yes
Trend*State Dummies	No	No	Yes	Yes
Transfers, Expenditure, Agric. Prod.	No	No	No	Yes
Observations	7,301	7,117	7,117	7,115
Pseudo R2	0.87	0.88	0.92	0.92

Note: (***) Significant at 1%; (**) significant at 5%; (*) significant at 10%. Marginal effects calculated at the sample means. Clustered standard errors at the municipality level in brackets. All the models have lagged activity, number of public and private competitors and municipality payroll. Subsample $1 \leq n^{pub} \leq 3$ includes all municipalities that had at least one and at most three public players in every period.

Appendix 3: CCPs

Table 4.11: CCP Logit for Public Players (BB and CEF)

	(I)	(II)	(I)	(II)
	Sample:1988-2010		Sample:1996-2010	
Lagged Activity	6.75***	6.73***	6.88***	7.22***
	[0.09]	[0.10]	[0.12]	[0.14]
Market Payroll	0.03***	0.02***	0.04***	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]
Sample Payroll		0.07***		0.17***
		[0.01]		[0.01]
Market 1		-1.15***		-1.17***
		[0.26]		[0.32]
Market 2		-1.09***		-1.45***
		[0.23]		[0.29]
Market 3		-0.64***		-0.88***
		[0.22]		[0.26]
Dummy BB	1.35***	1.68***	1.82***	2.22***
	[0.11]	[0.13]	[0.14]	[0.17]
Constant	-3.70***	-3.56***	-4.55***	-5.58***
	[0.12]	[0.25]	[0.16]	[0.33]
Observations	20,357	20,357	13,680	13,680
Pseudo R2	0.793	0.796	0.815	0.827

Note: (***) Significant at 1%; (**) significant at 5%; (*) significant at 10%. Clustered standard errors in brackets. Model I does not include sample payroll and market dummies. Model II includes these variables.

Table 4.12: CCP Logit for Private Players (Bradesco and Itau)

	(I)	(II)	(I)	(II)
	Sample:1988-2010		Sample:1996-2010	
Lagged Activity	7.44***	7.32***	8.05***	8.17***
	[0.12]	[0.13]	[0.18]	[0.20]
N Public	0.25***	0.25***	0.46***	0.63***
	[0.05]	[0.07]	[0.06]	[0.08]
N Private	-0.42***	-0.40***	-0.57***	-0.35**
	[0.10]	[0.11]	[0.12]	[0.15]
Market Payroll	0.02***	0.01***	0.01***	0.01***
	[0.00]	[0.00]	[0.00]	[0.00]
Sample Payroll		0.08***		0.03**
		[0.01]		[0.01]
Market 1		-0.42		0.98**
		[0.32]		[0.38]
Market 2		-0.12		0.85***
		[0.26]		[0.32]
Market 3		0.13		0.51**
		[0.21]		[0.26]
Dummy Bradesco	0.05	0.04	-0.51***	-0.52***
	[0.11]	[0.11]	[0.13]	[0.13]
Constant	-3.84***	-4.43***	-3.41***	-4.84***
	[0.11]	[0.39]	[0.13]	[0.46]
Observations	15,919	15,919	10,595	10,595
Pseudo R2	0.828	0.831	0.830	0.831

Note: (***) Significant at 1%; (**) significant at 5%; (*) significant at 10%. Clustered standard errors in brackets. Model I does not include sample payroll and market dummies. Model II includes these variables.

Chapter 5

Conclusions

The thesis investigates Industrial Organization questions. The first essay (chapter 2) uses a static framework to analyze market effects of Brazilian Biodiesel regulation. In the second essay (chapter 3) we propose an estimator for dynamic games. In the third essay (chapter 4) we use the estimator developed in the second paper to estimate an dynamic oligopoly model for the banking industry in Brazil.

To the best of my knowledge, no written work has used the data set I use in the first essay to analyze the welfare impacts of the biodiesel regulation. The results obtained are important as they show that the positive gains of the biodiesel producers are offset by the loss of surplus from consumers, wholesalers and retailers. These results, however, should be taken with caution. First, due to data constraints it is not possible to study the dynamics of this industry. Therefore, the results obtained are supposed to hold only in the short run. Second, the paper does not consider the effects of the regulation in the other part of the production chain. Small farmers, for example, may have welfare gains due to implementation of the biodiesel regulation. Finally, the environmental benefits are also not computed. Future work should deal with this questions, specially regarding the dynamic nature of this industry.

The estimator proposed in chapter 3 can be very useful in applied work. Monte Carlo experiments show a significant reduction in the computation time. The gains are specially important when the payoffs are linear in parameters, the practical leading case. In this case our estimator assumes a familiar OLS / GLS form. The estimator, however, cannot accommodate unobserved heterogeneity and suffers with small sample bias, two common problems in a two step estimator.

Chapter 4 has an important contribution as it uses a structural model to predict the effects of privatization policies. To the best of my knowledge, there are no papers using a similar approach to answer this kind of question. However, even after controlling for all observable variables, there may be some unobservable heterogeneity biasing the results. Future work should look for ways to control for these unobservables.