RICE UNIVERSITY

Four Essays on Applied Energy Economics and Policy

by

Raúl Bajo Buenestado

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE
Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Peter R. Hartley, Chair
George and Cynthia Mitchell Chair in Sustainable Development and Environmental Economics and Professor of Economics

Kenneth B. Medlock, III
James A. Baker, III, and Susan G. Baker Fellow in Energy and Resource Economics, James A. Baker III Institute for Public Policy Adjunct Professor of Economics

Xun Tang
Associate Professor of Economics

Leonardo Dueñas-Osorio
Associate Professor of Civil and Environmental Engineering

Houston, Texas
April, 2016
ABSTRACT

Four Essays on Applied Energy Economics and Policy

by

Raúl Bajo Buenestado

This thesis is divided in two parts. The first part (chapters 1 and 2) studies capacity payments in the electricity sector. The second part (chapter 3 and 4) is on gasoline retail markets.

The first chapter explores welfare implications of capacity markets in the electricity sector. We propose a theoretical model with cost heterogeneous firms, for which price and quantity equilibria are obtained both with and without a capacity market. The consequences for consumers are assessed using three different measures: consumer surplus, probability of blackout and price volatility. We conclude that a capacity market is able to reduce extreme events. Under some circumstances, we show that a capacity market is also efficiency enhancing.

In the second chapter, we use data from the Texas ERCOT to study the impact of capacity payments in a stylized wholesale electricity market. We find that the introduction of capacity payments has two countervailing effects. On the one hand, it increases consumers’ bills. On the other hand, it reduces price volatility and blackout probability. We find that the net impact on consumer surplus is negative both in a perfectly competitive market and in the presence of market power.

In the third chapter, we use monthly data from the Spanish gasoline retail market to explore asymmetries in consumers’ responses to changes in gasoline prices and taxes. We investigate whether an increase in taxes has a more negative impact on the
demand than an increase in the “pre-tax” price of gasoline. We estimate consumers’ behavioral responses using a rich set of robust models. We find evidence of asymmetric responses for the demand of unleaded fuels and agricultural diesel fuel.

In the final chapter we study a game of spatial competition in prices. We focus on the linear city duopoly model to see what we can learn about the distribution of consumers, which is approximated using variation in equilibrium prices and costs. We apply our methodology to a dataset on prices of a pair of gas stations in a straight highway. Using our approximation, we are able to calculate where should be located an entrant gas station to maximize welfare.
ACKNOWLEDGEMENTS

Forthcoming...
Contents

Abstract ................................................................. ii

List of Figures .......................................................... ix

List of Tables ............................................................. xi

1 Thesis motivation ......................................................... 1

I Power Markets .......................................................... 4

2 Regulated capacity and capacity payments in liberalized and competitive electricity sectors ............................................. 5

2.1 Introduction .......................................................... 5

2.1.1 Related literature ............................................... 9

2.1.2 Model overview ................................................. 12

2.2 Theoretical framework ............................................... 13

2.3 Equilibrium analysis ............................................... 17

2.3.1 Equilibrium in the wholesale spot electricity market ........ 17

2.3.2 Equilibrium capacities ....................................... 19

2.4 Welfare analysis ................................................... 22

2.4.1 Main results .................................................. 25

2.5 Conclusions ......................................................... 27
### Welfare consequences of price caps and capacity payments in the Texas electricity market (ERCOT)

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>29</td>
</tr>
<tr>
<td>3.2</td>
<td>Theoretical framework</td>
<td>35</td>
</tr>
<tr>
<td>3.3</td>
<td>Equilibrium in the wholesale electricity market</td>
<td>37</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Perfect competition in the wholesale electricity market</td>
<td>40</td>
</tr>
<tr>
<td>3.4</td>
<td>Equilibrium capacities</td>
<td>42</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Equilibrium capacities in a perfectly competitive market</td>
<td>44</td>
</tr>
<tr>
<td>3.5</td>
<td>Market power in electricity markets: dominant firm-competitive fringe model</td>
<td>47</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Equilibrium in the wholesale electricity market</td>
<td>49</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Equilibrium capacities</td>
<td>50</td>
</tr>
<tr>
<td>3.6</td>
<td>Welfare analysis</td>
<td>53</td>
</tr>
<tr>
<td>3.7</td>
<td>A case study of the Texas ERCOT market</td>
<td>55</td>
</tr>
<tr>
<td>3.7.1</td>
<td>Background</td>
<td>55</td>
</tr>
<tr>
<td>3.7.2</td>
<td>Data</td>
<td>58</td>
</tr>
<tr>
<td>3.7.3</td>
<td>Numerical analysis and results</td>
<td>61</td>
</tr>
<tr>
<td>3.8</td>
<td>Conclusion</td>
<td>66</td>
</tr>
</tbody>
</table>

### II Retail Gasoline Markets

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>69</td>
</tr>
<tr>
<td>4.2</td>
<td>Methods and data</td>
<td>77</td>
</tr>
<tr>
<td>4.2.1</td>
<td>The baseline model</td>
<td>77</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Robustness check I: partial adjustment models in gasoline demand</td>
<td>80</td>
</tr>
</tbody>
</table>
4.2.3 Robustness check II: lagged effect of unemployment on diesel B consumption ........................................... 81
4.2.4 Robustness check III: instrumental variable regression ................................................................. 81
4.2.5 Data and summary statistics ........................................................................................................ 82
4.3 Results and discussion .................................................................................................................. 86
  4.3.1 The baseline model .................................................................................................................. 86
  4.3.2 Partial adjustment models in gasoline demand ........................................................................... 91
  4.3.3 Lagged effect of unemployment on diesel B consumption ....................................................... 93
  4.3.4 Instrumental variable regression ............................................................................................... 95
4.4 Conclusions and policy implications ............................................................................................ 98

5 Where to locate a gas station? Socially optimal location in a linear city game with non-uniform distribution of consumers ......................................................... 100
  5.1 Introduction .................................................................................................................................. 100
    5.1.1 Related literature ......................................................................................................................... 104
  5.2 Game theoretic framework ......................................................................................................... 106
    5.2.1 Further theoretical aspects: entrant’s location and consumers’ welfare ...................................... 110
  5.3 Model application .......................................................................................................................... 112
    5.3.1 Approximating the distribution of consumers under known transportation cost ................. 114
    5.3.2 Parametric restrictions ............................................................................................................... 115
  5.4 The Spanish retail gasoline market data ..................................................................................... 119
    5.4.1 Main features of the market ...................................................................................................... 120
    5.4.2 Data and descriptive statistics ................................................................................................ 122
    5.4.3 Main results and counterfactuals ............................................................................................... 125
  5.5 Conclusions ..................................................................................................................................... 129
6 Thesis conclusions .................................................. 131

Bibliography .............................................................. 133

Appendices ................................................................. 148

A ................................................................. 149

B ................................................................. 158

C ................................................................. 165

D ................................................................. 169

E ................................................................. 170

F ................................................................. 173

G ................................................................. 176
List of Figures

2.1 Potential scenarios in the spot electricity market given capacity choices 16

3.1 ERCOT forecasted reserve margin (2011) 57

3.2 ERCOT hourly load duration curve, net of wind and nuclear generation (1996-2014) 61

5.1 Total transportation costs for different sections of the linear city 111

5.2 Map of the section of the AP-7 Toll Highway 124

5.3 Estimated distribution of consumers 126

5.4 Map of the section of the AP-7 Toll Highway 127

5.5 Incumbents’ locations and entrant’s socially optimal location 127

A.1 Cdf associated to $L$ (red) and cdf associated to $L^m$ (green) 155

E.1 Equilibrium prices for different firms’ locations, distributions of consumers and transportation costs 171

F.1 True and fitted distribution of the demand (true follows a beta) 174

F.2 True and fitted distribution of the demand (true follows a quadratic) 175

G.1 Base case: symmetric distribution and locations ($\tau = 25$) 176

G.2 Symmetric distribution and locations, lower transportation cost ($\tau = 24$) 176

G.3 Demand advantage for firm 1: left skewed distribution, symmetric location ($\tau = 25$) 177
G.4 Location advantage for firm 1, symmetric distribution ($\tau = 25$) . . . . 177
List of Tables

3.1 ERCOT System-Wide Offer Cap (SW-CAP) ......................... 58
3.2 Welfare gains of raising/eliminating the price cap (perfect competition) 62
3.3 Welfare gains of introducing capacity payments (perfect competition) 64
3.4 Welfare gains of raising/eliminating the price cap (imperfect competition) 65
3.5 Welfare gains of introducing capacity payments (imperfect competition) 66

4.1 Standard deviation of the main variables ($I = 50$, $T =$Jan’11-Oct’14) 85
4.2 Effects of prices and taxes on the consumption of unleaded fuels . . . 87
4.3 Effects of prices and taxes on the consumption diesel fuels . . . . . 89
4.4 Partial adjustment model for unleaded fuels ........................ 91
4.5 Partial adjustment model for diesel fuels .......................... 93
4.6 Model with lagged agricultural unemployment for diesel B ........... 94
4.7 IV regression for unleaded and diesel fuels .......................... 96

5.1 Transportation cost and shape parameters for different Beta distributions ($n = 2000$) .................................................. 119
5.2 The “highway premium”: evidence of higher prices in highways . . 121
5.3 Summary statistics ....................................................... 125
5.4 Percent of welfare gains from optimal location of gas stations vs. geographical mid-point locations .............................. 128

C.1 Summary statistics ....................................................... 166
C.2 Summary statistics for (nominal) tax data (by Autonomous Community) 167
Chapter 1

Thesis motivation

Currently, the energy sector has a crucial role in both advanced and developing countries. Even though the actual contribution of the energy sector to the GDP is not relatively large (especially in countries with no energy resources in the ground), a collapse in the electricity system, a sudden and extraordinary increase in gasoline prices or natural gas delivery problems, to name a few, could have an enormous impact on a country’s GDP. Indeed, stable and reliable energy access not only contributes to economic growth, but also is a key element in the provision of many other vital services, such as water and sanitation, public safety and healthcare in a modern economy.

In this context, energy policy becomes a crucial element of academic and government discussion. According to the International Energy Agency (IEA), energy policy involves a great array of desiderata apart from economic growth, including energy security, economic development, and environmental protection. Given the multidisciplinary nature of these desiderata, energy policy typically covers multiple disciplines, including Economics, Engineering, Political Science, Environmental sciences, etc. and their combination.

While acknowledging the importance of these alternative perspectives on energy policy, this thesis nevertheless narrowly focuses on economic issues. More specifically,
I use well-known economic techniques (namely game theory and panel data methods) to enrich our understanding of economic regulation of two energy markets—electricity and transportation fuels. Our goal is to provide some ideas to increase economic efficiency (welfare) and, ultimately, to provide some useful policy recommendations toward the improvement of the status quo.

In particular, I focus on issues related to energy security and energy access in two markets that are of an increasing (and convergent) interest: electricity markets and fuel markets. In chapters 1 and 2, I address the current debate on the desirability of capacity markets in the electricity sector. Chapter 3 and 4 touch on some key policy elements regarding retail fuel markets.

The debate on capacity markets is ultimately a discussion on energy access and security of supply (reliability). It has been particularly intense in regions like Texas, and countries like Germany, which are considering their potential implementation.

Chapter 1 is mostly descriptive and theoretical. In this chapter, I prove some general and theoretical propositions applying to a stylized competitive electricity market. The main contribution of this chapter is to characterize a condition under which a capacity market is welfare-enhancing. The key insight is that a capacity market acts as an “insurance mechanism” for consumers. That is, it eliminates the negative consequences of some extreme events, such as quantitative rationing of electricity and price-spikes, at the cost of making consumers to pay a “premium” on average, which is the capacity payment.

It is one thing to show that a proposition holds in principle. It is something else to say that it matters in practice. In order to judge the quantitative significance of our results we need to examine a realistic example. Thus, chapter 2 applies the insights from the previous theoretical to a particular case-study: the Texas ERCOT market. I evaluate my equilibrium market conditions by using some data on costs and demand from Texas ERCOT. In addition, and realizing that the Texas market may
not be perfectly competitive due to the presence of two dominant firms, we extend the original model by allowing some degree of market power. Then, we perform a number of numerical analyses of the effects of including or excluding a capacity market under different competitive conditions.

Chapter 3 analyzes the influence of taxes on the demand for fuels. Following some recent relevant contributions, we use a rich set of robust panel data models to evaluate the impact of changes in taxes and “pre-tax” prices on gasoline demand using a detailed dataset from Spain. By understanding consumers’ behavioral responses (typically called elasticities in economics) we are able to provide some tax policy recommendations for the particular case of the Spanish market.

Finally, chapter 4 addresses the question of energy access in the context of a regulated market for transportation fuels. Using very particular assumptions, we show how a duopoly fuel market can reveal information about the distribution of consumer demand. The latter can then be compared to direct evidence on the spatial distribution of traffic on the road. It therefore provides an indirect test of the validity of a widely-used theoretical model in transportation economics. This model has otherwise been difficult to test. We then show how the model, once validated, can be used to inform policy about the optimal location of new fuel stations on the highway.
Part I

Power Markets
Chapter 2

Regulated capacity and capacity payments in liberalized and competitive electricity sectors

2.1 Introduction

Many electricity markets around the advanced world have undergone (and are still undergoing) deep, major and intensive reforms. Before the nineties, vertically integrated electric utilities typically carried out the main duties of power supply, i.e. generation, transmission, distribution, and retail supply. Thus, the sector was dominated by large monopolies that supplied electricity to residential, commercial and industrial consumers within a defined geographic area. These monopolies were government-owned or, if private, usually subject to price and entry regulation –see Joskow (2008a).

In late eighties, and especially during the nineties, many advanced countries initiated a process of privatization and liberalization of the electricity sector. The first country to liberalize the market was the UK, with the introduction of the Electricity
Although each country introduced its own idiosyncratic features, there are some common elements that most of the countries have been trying to pursue. Thus, according to Joskow (2008a), the “textbook” architecture of desirable features for restructuring, regulatory reform and the development of competitive markets for power involves (at least) the following key components: the privatization of state-owned electricity monopolies; a vertical separation of potentially competitive segments (e.g. generation, marketing and retail supply) from segments that will continue to be regulated (distribution, transmission, system operations); the creation of a maximally feasible number of competing generators; the horizontal integration of operation of and investment in transmission facilities and network operations with the designation of a single operator managing the network, and associated regulation of network institutions; and the creation of a voluntary public wholesale spot energy market, in which a group of private and competitive power generators (the supply) meet another group of private and competitive electricity retailers (the demand) with the latter supplying the electricity to their customers.

A well designed electricity market also has to ensure long-term supply reliability\(^2\). In other words, supply-side market participants must guarantee that at some point in the future (except in extraordinary and unforeseen circumstances) there will be enough generation capacity to provide electricity to customers, even during peak periods of demand. Since it takes time to build new capacity, future demand has to be anticipated, and the increasing economic activity and population growth require higher available capacity in the market every year to meet a growing demand. However, new rules and acts in many different countries are encouraging a substitu-

\(^1\) Although Chile is the first country that introduced major privatization and competition reforms back in 1982, as Fischer et al. (2000) point out, for many years the main generator, distributor and transmission company were under common ownership.

\(^2\) The British Electricity Act 1989, in section 3A under the title “The principal objective and general duties of the Secretary of State and the Authority”, states that one of the goals for the Secretary of State and the Authority is “to secure a diverse and viable long-term energy supply”.
tion from dirtier methods to generate electricity –especially coal– into greener ones –typically, renewables– that tend to be less reliable and offer a less predictable supply.\(^3\) Thus, not only demand growth but also social and ecological goals require electricity generators to make further investments in new and better plants.

In liberalized markets, electricity generators are typically private firms. Thus, in order to make any new investment, they calculate –as would a private firm in any other industry– whether the revenue they expect to obtain by selling electricity can cover the cost of building a new facility or plant, including a competitive return on the capital invested. The facility will be built only if the returns are adequate given the risks involved. If a plant is built but subsequent returns are insufficient to yield a competitive rate of return, we will encounter the so-called revenue adequacy problem or “missing money” problem. More formally, Joskow (2013) states that the “missing money” problem arises when “the expected net revenues from sales of energy and ancillary services at market prices provide inadequate incentives for merchant investors in new generating capacity”.

A key explanation for the emergence of the “missing money” problem in many markets is the existence of price caps. Price caps –or ceiling prices– are administrative actions imposed in the spot electricity market limiting the maximum spot market price of electricity during scarcity periods. They are present in the vast majority of electricity markets. It is usually argued that they have been put in place to reduce market power, although a major purpose in practice may be to limit politically unpopular price-spikes.\(^4\) Regardless of the motivation, according to Hogan (2005), caps on prices in restructured and liberalized markets have reduced the payments to electricity generators –especially to the peaking plants– that could be used to fund investment in new plants or simply to cover their operating costs. This has, in turn,

\(^3\)For instance, recently the US Environmental Protection Agency (EPA) announced the Clean Power Plan Proposed Rule. The goal is to cut carbon emissions from power plants.

lead to capacity shortages.

Warned about the potential problem created by price caps and, in general, in order to solve the “missing money” problem, many electricity market operators (or their government overseers) have recently implemented formal capacity markets. In these markets, electricity generators compete to get revenues, independent of payments for energy supplied or ancillary services, that allow them to build new generation capacity that consumers –typically through electricity retailers– may require at some point in the future. This competition between generators to supply future capacity commonly takes place through bidding in centralized auctions but it may also result from decentralized bilateral agreements between generators and electricity retailers. With the funds they obtain, electricity generators are able to make new investments and guarantee a certain level of installed capacity, i.e. guarantee that the availability of resources is adequate to meet future demand. Thus, a capacity market intends to protect the consumers in the sense of avoiding inadequate capacity and hence blackouts or brownouts in future periods when price caps might otherwise result in insufficient capacity to cope with scarcity events. All these benefits come at the cost of making consumers pay a capacity compensation that will make electricity prices more expensive (but less volatile) on average.

Capacity markets are present nowadays in many countries and regions. This is true for instance in the PJM, NYISO, ISO-NE, Italy and Western Australia. But at the same time, there are some other markets, such as the Texas ERCOT, that do not rely on this kind of market to supply capacity. Similarly, outside the US, the electricity market in Alberta, Australia’s NEM and the market in Scandinavia have not yet implemented capacity markets. Instead, they rely on other kinds of mechanisms to guarantee that there is enough capacity to meet electricity demand. The discussion about capacity markets is far from being over, as a lot of open issues

\textsuperscript{5}As Hogan (2013) points out, the product of these markets is installed capacity, not energy.
remain unresolved. In fact, at the time of writing, there is a huge debate about the desirability of such markets in Germany, Texas ERCOT, Peru\textsuperscript{6} and Netherlands, among others.

### 2.1.1 Related literature

Even though the debate about capacity markets is hot and broad, the economic literature examining capacity markets is very scarce and limited in scope. Several authors have previously studied investment incentives in electricity markets focusing on the role of price caps—for instance Fabra et al. (2011), Zöttl (2011). However, to our knowledge, there are just a few papers addressing the potential role of capacity markets as a way to solve the underinvestment problem from a formal perspective, namely Brown (2014), Elberg and Kranz (2013), Schäfer and Schulten (2014), Crampes and Creti (2005) and Creti and Fabra (2007).

Brown (2014) formalizes and studies optimal capacity payments and the ability of them to solve the “resource adequacy problem”. He also derives the potential consequences of these payments in terms of prices, firms’ profits and market competition. He finds that these payments are able to mitigate the problem of underinvestment, at the cost of increasing market concentration and/or firms’ rents. However, we have several concerns about the assumptions that he uses in his model. First, he departs from a symmetric duopoly structure in the market\textsuperscript{7} without assuming the usually imposed heterogeneity in firms’ costs, both in production and capacity building.\textsuperscript{8} Second, he assumes that demand for capacity is a random variable and that electricity generators choose their capacity limits before observing it. In reality, in most of the current capacity markets—especially those using centralized auctions—the demand is \textit{ex-ante}.

\textsuperscript{6}See Harbord and Pagnonzi (2014).
\textsuperscript{7}Which \textit{per se} is a source of market power and higher prices, as we have learned since the nineties from the British experience.
\textsuperscript{8}As he himself recognized, “[f]uture research should account for asymmetries in firms’ costs of electricity generation and capacity investment” (Brown (2014), p. 111).
known (with some minimal variation) by electricity generators. On the other hand, he analyses welfare including profits for the firms. We explicitly avoid it, as we just want to focus on consumers’ welfare.\footnote{As Abito (2012) does in a similar context, and considering that we study a competitive-market framework, we focus our analysis just on consumers’ welfare. In the context of a competitive market, since expected producers’ profit is set to zero, producers’ welfare loses relevance.}

Another study that formally includes capacity markets is Elberg and Kranz (2013). They analyze the effect of a capacity compensation mechanism on the structure of the electricity market. In their theoretical set up, they assume a competitive fringe of electricity generators and a strategic firm, which maximizes profits –eventually by withholding capacity. We assume instead that there are two kinds of electricity generators with heterogeneity in costs, so we are able to capture the merit order. In addition, while they focus on market structure, our analysis focuses on welfare implications for consumers, as it is the main issue of the current debate in many countries.

As Crampes and Creti (2005), we assume a two-stage framework.\footnote{A two-stage framework, including an auction in the second step, is a usual assumption in stylized model that study the behaviour of electricity firms. In fact, Hortacsu and Puller (2008) empirically demonstrate (for the particular case of the Texas ERCOT market) that firms’ bidding behaviour is consistent with the predictions of the game-theoretic equilibrium. Besides, many other authors employ a similar two-stage framework, including Tishler et al. (2008), Milstein and Tishler (2012) and Fabra et al. (2011), among many others.} In the second stage of both their game and ours, electricity generators compete in a uniform-price auction to supply power to consumers. However, in the first stage they analyze potential capacity withholding by generators given some installed capacity for a known (forecasted) demand. By contrast, we solve for the actual installed capacity in equilibrium, provided that the demand is (ex-ante) unknown. Furthermore, while they assume exogenously imposed asymmetries in installed capacities we solve for them. Thus, Crampes and Creti (2005)’s paper is not a model of strategic investment but rather a model of strategic withholding. In addition, they do not provide an explicit analysis in terms of welfare.
The work by Creti and Fabra (2007) focuses on the short-run question. In particular, they address the question of whether a capacity market is able improve the balance between demand and supply in such a way that the installed capacity is able to meet the load. They leave the question of capacity adequacy –which is the center of our chapter– unaddressed. Although they also study the effect on welfare of introducing such a capacity market, they do not explicitly include firm cost heterogeneity, thereby ignoring the merit order.

Finally, Schäfer and Schulten (2014) present a particularly rich model. However, they focus on designing a capacity market in such a way that is able to replace old generators by renewable ones. Thus, their conclusions are restricted to the case in which this is the main goal.

A more abundant literature has examined capacity markets from an atheoretical perspective. For instance, using some numerical examples, Joskow (2008b) states that while reforms in the “energy-only” market can partially solve the “missing money” problem, capacity compensation mechanisms are necessary to fully mitigate such a problem since investment incentives would be inadequate without them. Along the same lines, Ausubel and Cramton (2010) focus the discussion on the recent experience in Colombia –although they also discuss the New England and California cases– to argue that capacity markets are able to mitigate risk, reduce market power in the spot market, and solve the underinvestment problem by coordinating new investment and helping to generate more transparent power markets. On the other hand, Harbord and Pagnozzi (2014) are more skeptical about the Colombia and the New England capacity markets. They criticize the use of descending clock auctions, alleging that they facilitate an increase in market power. As an alternative, they state that a sealed-bid auction is a more appropriate mechanism to allocate capacity. Cramton and Stoft (2005) defend the necessity of implementing capacity payments in the current

---

11The commonly employed term “energy-only” market captures the situation in which the only mechanism used in the power sector is a wholesale market, with no capacity payments.
restructured electricity market. They also propose an appropriate design of such payments in order to mitigate problems related to market power, withholding supply and risk.

Some other authors are not convinced about the potential benefits of capacity markets. For instance, Hogan (2013) states that there is little or no connection between capacity markets and real time spot market operations. Therefore, he argues that the “putative product” of capacity markets is hard to price and measure, given how separated it is from the actual energy product. As an alternative, he proposes the so-called operating reserve demand curve mechanism to solve the resource adequacy problem.\textsuperscript{12} Finally, Kleit and Michaels (2013a) are also against capacity markets, and in particular against its implementation in the Texas ERCOT market.

As a final comment on the atheoretical papers, we note that they all provide useful and new ideas, and contribute to the current policy debate around capacity markets. However, since none of them use theoretical models as a building block of their discussion, the conclusions that we can draw from them are limited and difficult to generalize.

\subsection*{2.1.2 Model overview}

This chapter provides a formal model that can be used to study capacity markets and their impact on consumers’ welfare. Following Tishler et al. (2008) and Milstein and Tishler (2012), we propose a two-stage game environment in which two kinds of electricity generators – base load and peak load – invest in capacity in the first stage (before demand is realized) and then compete in a wholesale spot market to produce and sell electricity to consumers.

Applying backward induction, we first deal with the second stage. Building on the seminal work by Fabra et al. (2006), we propose a uniform-price auction as the

\textsuperscript{12}Actually, in June 2014, the Texas ERCOT implemented an operating reserve demand curve to solve the underinvestment problem.
mechanism to allocate electricity in the spot market. In the first stage, electricity generators solve for the equilibrium investment in capacity by considering the expected profits we assess in the second stage. We study the first stage both in the absence and in the presence of a capacity market. Then, we analyze and compare consumers’ welfare under both scenarios.

We find that a capacity market is able to reduce both the blackout probability and the price volatility for the consumers, at the cost of reducing their average consumer surplus. However, we also show that under some circumstances, a capacity market is able to increase overall expected consumer surplus, yielding the efficient market outcome.\(^\text{13}\)

Notice that this chapter intends to provide some general (theoretical) conclusions, so an actual quantitative analysis is not provided. In Chapter 2 we take this theoretical conditions into action by employing data from the Texas ERCOT market.

The rest of the chapter is organized as follows. Section 2.2 provides the theoretical framework—including the main assumptions—that we use throughout the chapter. Section 2.3 characterizes first the equilibrium in the wholesale market, and then the equilibrium capacities (both in the absence and in the presence of a capacity market). Section 2.4 contains the welfare analysis. Section 2.5 concludes. Appendix A includes the proofs for all the propositions and corollaries we state throughout this chapter.

### 2.2 Theoretical framework

There is a unit measure of identical base load electricity generators (b) and a unit measure of identical peak load electricity generators (p).\(^\text{14}\) Both types of generators

---

\(^{13}\)In this thesis, we do not touch on the role of the reliability of power transmission lines. We recognize that this is an important caveat, since many blackout in real-world market are not related to generation, but to transmission issues. However, the transmission business is usually a heavily-regulated one (natural monopoly). This fact closes the door for an “economic market analysis”.

\(^{14}\)Usually, in actual electricity markets, there are more than two types of electricity generators with short-run costs (nuclear, hydropower, combined cycle, diesel, geothermal, etc.), as reflected in
invest in capacity, and produce electricity up to their capacity. They sell electricity to consumers\textsuperscript{15} in a wholesale spot market. In the basic theoretical framework, free-entry is assumed.

We assume that consumers’ demand takes the form of a rotated “L”. In particular, each consumer has a reservation price, which is technically called Value of Lost Load (VOLL). It reflects consumers’ opportunity cost of having an interruption of electricity supply during a blackout.\textsuperscript{16} For the sake of simplicity, we assume that there is unit one continuum of risk averse consumers, whose reservation price (VOLL) is denoted as $p^H > 0$, and whose associated utility function is $U(\cdot)$, with $U'(\cdot) > 0$ and $U''(\cdot) < 0$.\textsuperscript{17}

Consumers’ aggregate demand, $\theta$, is stochastic. In particular, we assume that $\theta$ is a non-negative random variable distributed according to some cumulative distribution function $F(\theta)$, defined for the support $[\underline{\theta}, \overline{\theta}]$. Without loss of generality, we normalize this support (i.e. the demand) such that $\underline{\theta} = 0$ and $\overline{\theta} = 1$. Furthermore, we assume that $F(\cdot)$ is strictly increasing and $F^{-1}(\cdot)$ exists. The timing in this economy is as follows:

1. Generators decide how much to invest in capacity, i.e. $k_i \geq 0$, $i \in \{b, p\}$.

2. Demand is realized, $\theta \in [\underline{\theta}, \overline{\theta}]$. Then, generators transform capacity into electricity and compete in a uniform-price auction to sell the electricity they produce to consumers in a centralized wholesale spot market.

By investing in capacity, generators incur a cost. In particular we assume that

\textsuperscript{15}In reality, a set of firms, called retailers, channel the electricity to consumers through a retail market.

\textsuperscript{16}For more information about VOLL, see Cramton et al. (2013), Spees et al. (2013). Alternatively, we can think that this is the price at which consumers/retailers are indifferent between buying electricity in the spot market and building a self-generating plant that provides electricity if a blackout occurs. The whole point is to emphasize that this price is very high!

\textsuperscript{17}We do not assume a specific utility function, since the relevant element in our analysis is not individuals’ demand but aggregate demand.
electricity generators face constant linear costs of building $k_i \geq 0$ units of capacity, denoted as $c_{k_i} > 0$, $i \in \{b, p\}$, and such that $c_{k_b} > c_{k_p}$.

Next, generators produce electricity out of capacity according to the following production function:

$$0 \leq q_i \leq k_i \quad i \in \{b, p\}$$

so in the simplest version generators transform one unit of capacity into one unit of electricity.

Electricity is traded in the spot market at price $p^s > 0$. Moreover, base load generators incur a variable cost of $c_b > 0$ for producing electricity out of capacity while for a peak load generator variable cost of production is $c_p > 0$. This assumption captures the usual so-called merit order observed in electricity markets.

Let us denote the total installed capacity in the market as $K$, where $K = k_b + k_p$.

Thus, once the demand $\theta$ is realized, we face three possible scenarios:

1. $k_b \geq \theta$. I.e. base load capacity is greater than aggregate demand. Thus, base load generators have enough capacity to serve the whole demand.

2. $k_b < \theta < K$. I.e. base load capacity is not enough to serve the demand, but the whole demand can be served by both types of generators.

3. $K \leq \theta$. I.e. there is not enough capacity to serve the whole demand.

---

18 This assumption is consistent with the fact that per-unit capacity cost for the peak load generators is usually smaller than the per unit capacity cost for the base load generators. See Hartley and Moran (2000) and Public Service Commission of Wisconsin (2012).

19 Since we assume in accordance with the actual rules governing electricity wholesale markets that the allocation mechanism in the spot market is a uniform-price auction, the equilibrium price is paid to all the electricity generators that sell electricity in the market. The usage of this kind of auctions in the spot electricity markets is standard in the literature and consistent with the vast majority of real-world markets –see Von der Fehr and Harbord (1998) and Newberry (2002). It is also the main object of study in Fabra et al. (2006), and employed by Brown (2014) in the study of optimal capacity compensation mechanisms. We use a similar setup as the aforementioned authors.

20 Since we ignore wind generation and (consequently) subsidies to wind generation in our stylized model, we rule out the possibility of negative prices in our setup.

21 To rule out uninteresting scenarios, we restrict to the cases in which $k_i \geq 0$ for some $i \in \{b, p\}$ and $K \leq 1$. 
Each scenario is plotted separately in Figure 2.1. Notice that in the latter case (subfigure 2.1c), since the installed capacity is not enough to serve the whole demand, part of the demand side – the grey area – suffers a “blackout”.\textsuperscript{22} As we formalize later, the (ex-ante) probability of a blackout depends on the capacity choices of both types of generators in the first period. This issue, together with some other consequences for consumers, are discussed in upcoming sections. As Creti and Fabra (2007) do, we proceed by backward induction. Thus we first calculate the equilibrium quantities and prices for both the base load generator and the peak load generator in the wholesale spot market given the capacity choices. Given these quantities and prices, we then obtain the equilibrium capacities.

Figure 2.1: Potential scenarios in the spot electricity market given capacity choices

\begin{itemize}
  \item[(a)] $k_b \geq \theta$
  \item[(b)] $k_b < \theta < K$
  \item[(c)] $K \leq \theta$
\end{itemize}

\textsuperscript{22}This may not literally be a blackout of the whole network but rather a curtailment of supply to customers with interruptible demand contracts.
2.3 Equilibrium analysis

In this section, we analyze both the equilibrium capacity choices in the first period as well as the equilibrium production and price in the wholesale spot electricity market (the second period). Since electricity generators choose capacity based on the expected profit that they make out of selling electricity in the spot market, we address first the latter market. Based on our results, we calculate the equilibrium capacity choices afterwards.

2.3.1 Equilibrium in the wholesale spot electricity market

Given the capacities that both type of generators choose in the first period, $k_i$, $i \in \{b,p\}$, these generators compete in a centralized uniform-price auction to supply electricity to consumers. In particular, once they observe the realization of the demand, each generator submits an offer price (bid), $p_i$, $i \in \{b,p\}$. Through this offer price (bid), a generator specifies the minimum price at which it is willing to produce and sell electricity. An independent agent (in practice an independent system operator, or ISO) clears the auction as follows. The lower bidding generator is dispatched (and thus serves the market) first. The higher-bidding generator is dispatched afterwards to serve the residual demand if capacity from the lower-bidder generator is not enough to cover the whole demand $\theta$. In case of a tie, the rationing rule is just prorata between marginal bidders.\(^{23}\)

Thus, for a given offer price profile $\vec{p} \equiv (p_b, p_p)$, the output allocated to type-$i$ generator, $i \in \{b,p\}$, denoted by $q_i(\theta; \vec{p})$ is:

\(^{23}\)In practice, factors such as different locations on the network, different transmission line capacities and different locations of demand would favor one generator over another.
\[ q_i(\theta; \vec{p}) = \begin{cases} 
\min\{\theta, k_i\} & \text{if } p_i < p_j \\
\frac{1}{2}\min\{\theta, k_i\} + \frac{1}{2}\max\{0, \theta - k_j\} & \text{if } p_i = p_j \\
\max\{0, \theta - k_j\} & \text{if } p_i > p_j 
\end{cases} \]

(2.2)

where \( j \in \{b, p\}, j \neq i \).

Since the quantity-allocation mechanism is a uniform-price auction, the equilibrium price in the spot market, \( p^s \), is equal to the higher accepted price offer (bid). Therefore, for each unit of electricity produced and sold, generator \( i \) receives \( p^s \). Given this environment, the following proposition solves for the price, production and (ex-post) wholesale aggregate profits in equilibrium.

**Proposition 1.** In equilibrium, given electricity generators’ capacities, \( k_i \) for \( i \in \{b, p\} \):

1. If \( k_b \geq \theta \) the spot market price is equal to base load generators’ variable cost. I.e. \( p^s = c_b \). Moreover, base load generators serve the entire demand \( (q_b = \theta) \) and generators earn zero profit \( (\pi_i^s = 0, i \in \{b, p\}) \).

2. If \( k_b < \theta \leq K \) the spot market price is equal to peak load generators’ variable cost. I.e. \( p^s = c_p \). Moreover, base load generators produce at maximum capacity \( (q_b = k_b) \) and peak load generators serve the residual demand \( (q_p = \theta - k_b) \). Base load generators’ aggregate profit is positive and such that \( \pi_b^s = c_p k_b - c_b k_b \) while peak load generators earn zero profit \( (\pi_p^s = 0) \).

3. If \( K < \theta \) the spot market price is equal to the Value of Lost Load (VOLL). That is, \( p^s = p^H \). Moreover, both base load generators and peak load generators produce at maximum capacity \( (q_i = k_i, i \in \{b, p\}) \). Aggregate profits are such that \( \pi_b^s = p^H k_b - c_b k_b \) and \( \pi_p^s = p^H k_p - c_p k_b \).

The intuition behind this result is the following. If base load capacity is greater than the demand \( (k_b \geq \theta) \), base load generators have enough capacity to serve the
whole demand. Since these generators are price-takers, competition drives the price to their marginal cost, $c_b.^{24}$ Hence $p^s = c_b$. If base load capacity is not enough to serve the demand ($k_b < \theta < k_b + k_p$) then there is room for peak load generators to produce and sell electricity. Again, since these generators are price-takers, competition drives the price to their marginal cost, $c_p$, and both types of generators receive the same $p^s$, which is given by $p^s = c_p$. In the third case, capacity is scarce. Hence (excess) demand drives the price up to $p^H$ (VOLL), which is the price consumers are willing to accept to suffer an interruption of power supply (the reservation price), assuming that we have customers with interruptable contracts. The part of the demand that is not covered suffers a power outage.

2.3.2 Equilibrium capacities

Next, we examine the problem that generators solve in the first period. At this stage, firms decide the amount of investment in capacity, $k_i$ for $i \in \{b, p\}$. This decision is taken before knowing the realization of demand. However, capacity levels are chosen anticipating period-2 equilibrium profits.

We restrict our analysis to the cases in which $k_i \in [0, 1]$. On the one hand, it does not make sense to talk about negative investment in capacity (divestment), since firms enter period 1 with zero investment in capacity.^{25} On the other hand, it also seems unreasonable for the firms to choose a level of capacity greater than 1, since investment is costly for them and part of their capacity would be unused.

We examine equilibrium aggregate capacities both with and without a capacity market. In the former case, the regulator establishes a capacity target level such that the aggregate demand is fully served in period 2, regardless of the realization of $\theta$. Firms compete to achieve such a regulated capacity target and they receive a capacity

---

^{24} Obviously, peak load generators do not produce and sell electricity in this case.

^{25} It could make sense to talk about negative capacity investment in a dynamic setting, for instance, if a firm loses demand over time.
compensation payment, which is determined through competition.\footnote{Alternatively, the regulator could make a regulated capacity compensation payment to firms competing to invest in capacity.}

Analytically, this mechanism is modeled as follows. The regulator fixes a capacity target level, $K^T$, that must be equal to the sum of the total capacity investment that both types of firms make in equilibrium.\footnote{Given our initial assumptions, in order to avoid shortages, the capacity target level must be $K^T = 1$, which implies $Pr(K^T \leq \theta) = 0$ (no-shortages case). In real-world markets, the adequate capacity level is calculated by the regulator, using the so-called rule of 1-hour-in-10-years.} In other words, $K^T = k^r,m_b + k^r,m_p$, where $k^r,m_i$ is the equilibrium aggregate capacity made by type-$i$ suppliers. To achieve this capacity level, the firms that invest in capacity receive a compensation $m$ per unit of capacity built, which is determined through competition, using the long-run free-entry equilibrium as our solution concept. Such a capacity compensation is passed-through to consumers.

**a) Scenario #1: equilibrium without a capacity market**

As we mention above, free-entry is assumed in the first stage. More precisely, we assume that the number of generators in period 1 is the “free-entry equilibrium” number of generators. That is, $n_i$ is such that $\mathbb{E}\pi_i \geq 0$ if the number of generators is $n_i$, and $\mathbb{E}\pi_i < 0$ if the number of firms is $n_i + r$, for $r = 1, 2, \cdots$, where $\mathbb{E}\pi_i$ represents (expected) aggregate period-2 profit for type-$i$ firms at period 1. Considering that the number of potential entrants is large and their size is infinitesimally small (unit measure of perfect competitors), we ignore the integer constraint on the number of firms. Thus, we assume that the free-entry equilibrium number of suppliers exactly satisfies the zero-profit condition, i.e. $\mathbb{E}\pi_i = 0$. Having said that, the following proposition captures the equilibrium aggregate capacities in the absence of a capacity market and capacity compensation payments.

**Proposition 2. Equilibrium capacities without a capacity market.** Denote $a_i \equiv p^H - c_i$, $i \in \{b, p\}$. In a perfectly competitive market at an interior so-
the unique equilibrium aggregate capacity of base load generators is $k^*_b = F^{-1}\left(1 - \frac{c_{kp}-c_{kb}}{c_{p}-c_{b}}\right)$, and the unique aggregate capacity of peak load generators is $k^*_p = F^{-1}\left(1 - \frac{c_{kp}}{a_{p}}\right) - F^{-1}\left(1 - \frac{c_{kb}}{c_{p}-c_{b}}\right)$.

From Proposition 2 it follows that in equilibrium, even if both types of generators choose a strictly positive amount of investment in capacity, the total installed capacity in the market is less than 1 (the maximum peak demand). In other words, in a competitive market, underinvestment always occurs in the sense that the market does not provide enough capacity to serve all the consumers in the highest possible realization of the aggregate demand. Intuitively, the capacity constraint must be binding with positive probability so peak load generators can earn the excess of price over the marginal cost in those states. Such “short-run rents” are needed in turn to pay for the investment in capacity.

**Corollary 1.** In equilibrium, the total installed capacity is strictly less than one. That is, $K^* \equiv k^*_b + k^*_p < 1$.

Corollary 1 implies that in the absence of regulated capacity and capacity payments, a competitive market does not achieve the investment level necessary to avoid shortages at some high levels of demand.

**b) Scenario #2: equilibrium with a capacity market**

As discussed above, in a regulated capacity regime, the regulator fixes a target capacity level $K^T$ that electricity generators in the market must achieve. To achieve this target level, the generators that invest in capacity receive a compensation $m$ per unit of capacity built. The next proposition solves for the compensation that firms

\footnote{We restrict our analysis to the interior solution case, i.e. to the case in which the investment in both types of technologies is strictly positive. The reason is that, as is the case in reality, we are interested in investigating a market in which a homogeneous good is produced with more than one technology. This fact does not imply, in accordance with reality, that when the centralized market clears all the homogeneous good is produced using only one technology. For the sake of completeness, we include the corner solutions in Appendix A.}
are willing to accept to achieve an equilibrium in which the level of investment is given by $k_i^{*,m}, i \in \{b, p\}$, and such that $K^T \equiv k_b^{*,m} + k_p^{*,m}$.

**Proposition 3. Equilibrium capacities with regulated capacity.** Denote $a_i \equiv p^H - c_i, i \in \{b, p\}$. Given $K^T$, in a perfectly competitive market at an interior solution the unique equilibrium aggregate capacity of base load generators is $k_b^{*,m} = F^{-1} \left( 1 - \frac{c_b - c_p}{c_p - c_b} \right)$, the unique equilibrium aggregate capacity of peak load generators is $k_p^{*,m} = K^T - F^{-1} \left( 1 - \frac{c_b - c_p}{c_p - c_b} \right)$, and the unique per-unit capacity equilibrium compensation payment is $m = c_k + a_p(K^T - 1)$.

Using this result, we can show that in the presence of regulated capacity and capacity compensation payments, the regulator solves the underinvestment problem by setting a sufficiently high target capacity level.

**Corollary 2.** There exists an $m > 0$ such that $K^T > K^*$.

However, this gain comes at a potential cost. In particular, in the next section we show that although a regulated capacity regime is able to solve the underinvestment problem, it may decrease consumers’ welfare. This is due to the fact that, as discussed above, the capacity compensation is effectively passed through to final consumers. Through the rest of the chapter we build upon the previous results (Propositions 1, 2 and 3) to analyze the welfare implications for consumers in a perfectly competitive market.

### 2.4 Welfare analysis

As mentioned above, our aim is to evaluate the desirability of capacity markets for consumers in liberalized and competitive electricity sectors. For that purpose, in this section we develop a consumers’ welfare analysis of the different proposed scenarios. We compare the no-capacity-market equilibrium with the equilibrium obtained in the presence of a capacity target and capacity compensation payments.
As is standard in the literature, we measure consumers’ welfare in terms of \((ex-ante)\) consumer surplus, i.e. the maximum price at which consumers are willing to buy electricity (VOLL) minus the price they pay multiplied by the total amount they purchase. In the absence of a capacity market, consumer surplus is given by:

\[
CS = \int_{k_b}^{k_b} (p^H - c_b) \theta dF(\theta) + \int_{k_b}^{K} (p^H - c_p) \theta dF(\theta) \tag{2.3}
\]

In the presence of a regulated capacity target level \((K^T)\) and a per-unit capacity compensation payment \((m)\), consumer surplus is given by:

\[
CS^m = \int_{k_b}^{k_b} (p^H - c_b) \theta dF(\theta) + \int_{k_b}^{K^T} (p^H - c_p) \theta dF(\theta) - m \tag{2.4}
\]

where \(m\) is the equilibrium compensation payment that is passed-through to consumers.

The consumer surplus expressions depend on two key components. First, they depend on the market price of electricity. In particular, an increase in the price of electricity decreases consumers’ welfare. Second, these expressions also depend on the equilibrium capacity levels. In particular, an increase in aggregate capacity reduces both the probability of excess demand (which drives the market-clearing price up to the reservation price) and the probability of shortage and, consequently, makes consumers \((ex-ante)\) better off.

In addition to consumer surplus, we also study price volatility (variance) and price-spike risk. For that purpose, let us define the set of market-clearing prices \(C \subseteq \mathbb{R}_+\). Denote \(n\) as the \((ex-ante)\) possible number of equilibria in the spot market in period 2, and consider the set of \(n\) market-clearing prices, \(C^n = \{p_1^s, \ldots, p_n^s\} \subseteq \mathbb{R}_+^n\). Given the space of market-clearing prices, we denote the set of lotteries over \(C^n\) as \(\mathcal{P} = \Delta(C^n)\) which is an \(n\)-dimensional unit simplex, i.e. \(\Delta(C^n) = \{\alpha_1, \ldots, \alpha_n | \alpha_i \geq 0 \forall i \in n \text{ and } \sum_i \alpha_i = 1}\). We capture both a set of market-clearing prices together
with a lottery associated to it in a price-contingent contract for consumers.

**Definition 1.** Given a set of \( n \) market-clearing prices, \( C^n \equiv \{p^s_1, \ldots, p^s_n\} \subseteq \mathbb{R}^n_+ \), where \( n \) is the (ex-ante) possible number of equilibria in the spot market in period 2, and given the set of lotteries over \( C^n \), \( \mathcal{P} = \Delta(C^n) \), such that \( \Delta(C^n) = \{\alpha_1, \ldots, \alpha_n \mid \alpha_i \geq 0 \forall i \in n \text{ and } \sum_i \alpha_i = 1\} \), we refer to \( \mathcal{L} \) as the price-contingent contract for consumers associated to \( C^n \), denoted as:

\[
\mathcal{L} = \begin{pmatrix}
p^s_1 & \cdots & p^s_n \\
pr(1) & \cdots & pr(n)
\end{pmatrix}
\]

and we denote \( G_{\mathcal{L}}(p^s) \) the cdf associated to the price-contingent contract \( \mathcal{L} \) such that \( G_{\mathcal{L}}(p^s) \equiv pr(p^s_r \leq p^s) \), \( \forall p^s \in C \), \( r \in \{1, \ldots, n\} \).

Next, to compare price riskiness between two price-contingent contracts, we employ the commonly used notion of Second Order Stochastic Dominance (SOSD):

**Definition 2.** Given two price-contingent contracts with the same mean, \( \mathcal{L} \) and \( \mathcal{L}' \), \( \mathcal{L} \) Second Order Stochastic Dominates (is less risky than) \( \mathcal{L}' \) if and only if \( \int_{-\infty}^{p^s} G_{\mathcal{L}}(z)dz \leq \int_{-\infty}^{p^s} G_{\mathcal{L}'}(z)dz \), \( \forall p^s \in C \).

As shown in Mas-Colell et al. (1995), SOSD implies that a price-contingent contract that is less risky than (i.e. that Second Order Stochastic Dominates) another is weekly preferred by all risk-averse consumers.

Even though the concept of SOSD is stronger than the concept of variance in the presence of risk-averse consumers, it requires both price-contingent contracts to have the same mean, which is a restrictive assumption. Therefore, we also study the variance of the price, as a measure of price volatility, which does not require such a restrictive assumption.

\[\text{Notice that the concept of SOSD is stronger than the concept of variance, as a measure of price volatility. In fact, we can show that } \mathcal{L} \text{ Second Order Stochastic Dominates (is less risky than) } \mathcal{L}' \iff \text{var} (\mathcal{L}) \leq \text{var} (\mathcal{L}') \text{, but } \text{var} (\mathcal{L}) \leq \text{var} (\mathcal{L}') \iff \text{that } \mathcal{L} \text{ Second Order Stochastic Dominates (is less risky than) } \mathcal{L}' \text{.}\]
2.4.1 Main results

Using the concepts mentioned above, we compare the no-capacity-market scenario with the capacity-market scenario. We focus our analysis on the case in which the regulated capacity target avoids shortages, i.e. \( Pr(K^T \leq \theta) = 0 \), solving the under-investment problem that occurs in equilibrium in the absence of regulated capacity—see Corollary 1.

We first deal with consumer surplus. The following theorem captures the conditions under which a capacity market improves consumer surplus relative to an unregulated capacity regime.

**Theorem 1.** In equilibrium, if \( p^H - c_p > \frac{c_{Kp}}{\int_K^{K^T} \theta dF(\theta)} \), the presence of a capacity market unambiguously increases consumer surplus in comparison to the case in which there is no capacity market.

Given firms’ costs, Theorem 1 implies that a capacity market improves consumer surplus under the following conditions. First, given \( p^H \), this result holds if \( \int_K^{K^T} \theta dF(\theta) \) is close to one. This is true for relatively right-skewed distributions, since, if so, the expected value of peak demand will be close to \( K^T \), i.e. it will be close to 1. Therefore, and considering that from the proof of Proposition 2 we know that \( p^H - c_p > c_{Kp} \), then \( CS^m > CS \). The intuition of this result is as follows. If the probability of peak demand is relatively high, a no-capacity-market regime (which leads to underinvestment in capacity in equilibrium) will produce blackouts with relatively high frequency, reducing consumer surplus as shown above. Therefore, in this case, it is consumer-surplus-enhancing to introduce a capacity target and capacity payments, which avoid shortages.

Second, even if the risk of peak demand is not high, Theorem 1 also holds if consumers’ reservation price, \( p^H \) (VOLL), is high enough. Notice that the LHS of the inequality in Theorem 1 increases linearly—on a one-to-one exchange basis—as \( p^H \)
increases. The RHS of the inequality also increases as $p^H$ grows. In particular, as $p^H$ increases, the lower integrand ($K$) approaches to $K^T$, so the denominator converges to zero, making the whole RHS converge to infinity. However, the rate of increase of the LHS is greater than the rate of convergence of the RHS, particularly if $F(\cdot)$ is not heavily left-tailed (the more right-skewed is $F(\cdot)$, the slower is the convergence). Therefore, even if the risk of peak demand is relatively large (right-skewed distribution), for a high enough consumers’ reservation price, a capacity market increases consumer surplus. The intuition is that, if the maximum price (VOLL) that consumers are willing to pay is high, shortages will be too costly for them.

Finally, we provide two major results regarding price riskiness. Our first result is stronger than the second one, although it requires stronger assumptions.

**Proposition 4.** In equilibrium, if $a_p(1 - K) \geq c_{kp}$, for $p^H$ high enough, the price-contingent contract in a scenario with a capacity market, is less risky than (SOSD) the price-contingent contract in a scenario with no capacity market.

In other words, if the expected price of both contracts is the same –for which $a_p(1 - K) = c_{kp}$ is a sufficient condition–, a capacity market not only reduces price-spike risk, but it also reduces price volatility (measured as the variance). Thus, a price-contingent contract in the presence of regulated capacity and capacity payments is weekly preferred by all consumers to a price-contingent contract in a market with unregulated capacity.

Proposition 4 presents a strong result in terms of price volatility. However, it requires strong assumptions. Thus, in the next proposition we present a result in terms of price volatility that does not require restrictive assumptions.

**Proposition 5.** In equilibrium, the price-contingent contract in a scenario with a capacity market is less volatile than the price-contingent contract in a scenario with

---

$^{30}$In particular, if $p^H \geq \frac{c_{kp}}{1 - K} + c_p$.

$^{31}$This condition holds, for instance, if $F(\cdot)$ is uniformly distributed.
no capacity market, in the sense of having less variance.

In summary, in comparison to the capacity market scenario, the presence of regulated capacity and capacity payments allows consumers to avoid both shortages and price-spike risk. Moreover, under some circumstances, a capacity market also increases (ex-ante) consumers surplus, namely if the frequency of peak demand is relatively high or if consumers’ reservation price is relatively high. If so, a capacity market is a Pareto improvement for consumers. Otherwise, consumer surplus will be greater in the absence of a capacity market.

2.5 Conclusions

The special features inherent in the power industry have required a high level of attention from market regulators all around the world. Perhaps due in part to these difficulties this industry was essentially controlled by the public sector before the nineties. The liberalization pattern that has prevailed since that period has proved to be a real challenge for these regulators. They have struggled to balance the tradeoff between allowing bona fide market freedom while regulating the idiosyncratic factors that are inherent to the electricity sector and that could potentially lead to market power and, thus, to a worse outcome for the consumers.32

Among recent measures implemented in restructured and liberalized electricity sectors, the introduction of capacity markets is among the most controversial. In fact, while several countries and regions have implemented some form of capacity compensation mechanism—for instance, PJM, Colombia, Italy and ISO-NE—some others have not implemented them yet—for instance, ERCOT, NEM or Alberta. Thus, whether capacity markets are beneficial or not in the electricity sector is a question of current debate. That debate has been especially intense in countries like

32As Michaels (2007) points out, “Electricity has been the last and most difficult of the great deregulations, thanks to technology, economics and politics”.

Germany and regions like Texas, which are considering their introduction.

Throughout this chapter, we have studied the implications for consumers of the implementation of capacity compensation payments. For that purpose, we have proposed a theoretical model with cost heterogeneous electricity generators that invest in capacity to produce and sell electricity in a wholesale spot market once they have built their facilities.

Putting aside the question of whether the capacity compensation payments should be allocated using one mechanism or another –e.g. via auction or via bilateral trading– which is beyond the scope of this chapter, we have priced this compensation in equilibrium, to conclude that a capacity market serves as an insurance mechanism for the consumers. In other words, the introduction of a capacity compensation mechanism increases the amount of investment and mitigates some of the harm associated with the underinvestment problem while enhancing the risk reducing benefits.

Under some circumstances –for instance, if the VOLL is very high– the introduction of a capacity markets also improves consumer surplus in addition to reducing price volatility and reducing the likelihood of constrained power supply. Thus, in this case, the capacity market leads to a Pareto improvement in the electricity sector.

With this chapter we have sought to enrich the current debate about capacity markets. We have focused on the consequences of capacity markets for final consumers of electricity. Thus, this model intends to be a building block for future and more complex analysis, which may extend the current simple model that we propose. Further extensions should take into account the fact that in some countries the electricity sector, far from being competitive, is controlled by a few firms that behave strategically. Furthermore, we also think that a valuable extension should take into account the dynamic setting of the model we solve. Finally, we could examine the consequences of operating reserve demand curve mechanisms, such as the one introduced in Texas in June 2014.
Chapter 3

Welfare consequences of price caps and capacity payments in the Texas electricity market (ERCOT)

3.1 Introduction

In the late eighties and especially during the nineties, many countries initiated a process of privatization and liberalization of the electricity sector. The goal was to create a wholesale electricity market with competition among private firms substituting for regulation as a means of constraining prices while encouraging innovation and cost control. Through this privatization process, policymakers wanted to get rid of the vertically integrated public monopolies that carried out generation, transmission, distribution and retailing of power supply. The basic idea was that the key natural monopoly element in the industry resided in the “wires” business and the operation of the network. Generation in particular was seen as a business with increasing short-run marginal costs that could be operated competitively while the monopoly elements were still subject to regulation.
However, due to special features inherent to the supply of electricity, this new competitive framework also faces conflicting goals. On the one hand, a modern and liberalized power sector must guarantee long-term supply reliability. In other words, supply-side market participants must ensure that –except in extraordinary and unforeseen circumstances– there will be enough generation capacity to provide electricity to customers, even during peaks of demand and reasonably foreseen outages in generating or transmission capacity. In liberalized markets, the task of building generation capacity is, by design, entrusted to private electricity generators. Hence, –as would a private firm in any industry– generators would decide to build additional capacity only if the revenue they expect to obtain by selling electricity in the market can cover the cost of building a new facility or plant, including a competitive return on the capital invested.

On the other hand, another goal of current liberalized markets is to provide strong incentives to control costs and prices by increasing competition among the generators that produce and sell electricity in the wholesale market. However, a more competitive environment might result in a lower rate of return for generators and, hence, decrease incentives to build additional capacity. In other words, even though a more competitive wholesale market may result in lower operating costs and lower prices for consumers, it may also generate a problem of underinvestment in capacity.

One might think that, even in a market that is reasonably competitive overall, a generator building a peaking plant –i.e., a plant with relatively large variable costs that is dispatched only when all other plants are operating at capacity– may face more limited competition and very inelastic demand and thus be able to charge a sufficiently high price whenever it produces and sells in the market. In turn, high prices at peak times would also accrue to all other generators supplying the market at that time since all active suppliers receive the same price at all times. However, in the vast majority of electricity markets, this is precluded by price caps. In other words,
there is a legally binding maximum price at which generators can sell electricity in the wholesale market. Some authors have argued that price caps are imposed to reduce market power, to reduce investment costs and to limit unpopular price-spikes –see Brown (2014), Cabral and Riordan (1991) and Joskow and Tirole (2007). However, as Hogan (2005) point out, price caps exacerbate the revenue adequacy problem or “missing money” problem\(^1\); i.e. they further discourage needed investment in generation capacity.

Perhaps after experiencing the problems created by price caps or, more generally, in order to solve the “missing money” problem, many countries and regions implemented capacity adequacy requirements and capacity compensation payments in a so-called capacity market. Through such a market, electricity generators compete for funds to build new generation capacity that consumers –typically through electricity retailers– may require at some point in the future. Thus, capacity markets have been recently implemented in the PJM, NYISO, ISO-NE, Italy, UK and Western Australia, among others. However, there are some other markets, such as Texas ERCOT or Germany, in which regulators are skeptical about their potential benefits. In light of the debate in these countries and regions, this chapter uses some of the theoretical conditions in the previous chapter to address the following question: how much do consumers gain by implementing capacity adequacy requirements and payments relative to an “energy-only” market? In particular, we use data from the Texas ERCOT market to study whether the introduction of capacity compensation payments could improve consumers’ welfare in Texas.

The debate around the desirability of capacity markets within the ERCOT region of Texas has been particularly intense in recent years. In May 2011 the North American Electric Reliability Corporation (NERC) projected that the existing installed capacity in ERCOT might be insufficient to cover the peak demand in the summer

\(^1\)See Joskow (2013).
of 2011. In fact, three months later (in August 2011), Texas suffered a heat wave that severely strained the supply. As the Energy Information Administration (EIA) pointed out, the under-forecast of peak loads together with an inadequate reserve margin resulted in controlled rolling blackouts in which up to 4,000 MW of load was shed. This event raised questions about the success of ERCOT’s privatization process, as the installed capacity was clearly insufficient and the goal of long-term supply reliability appeared to be in jeopardy.

To solve this revenue adequacy problem, ERCOT proposed two policy measures after 2011. First, it raised the price cap from $2,500/MWh in 2011 to $9,000/MWh in 2015; i.e. the regulator created incentives to build additional capacity through more generous scarcity-pricing. Second, in 2014, ERCOT implemented a system to compensate the firms holding reserve installed capacity during scarcity events.² But a capacity market is still far from being considered in ERCOT, as the regulator remains skeptical about its benefits. However, many authors have questioned this skepticism after observing that capacity markets were successful in stimulating investment in additional capacity in some other countries.³

Even though the debate about capacity markets is broad, the economic literature examining capacity markets is very scarce and limited in scope. Several previous authors have studied investment incentives in electricity markets and the role of price caps –for instance Crampes and Creti (2005), Creti and Fabra (2007), Fabra et al. (2011) and Zöttl (2011), among others. However, to our knowledge, there are just a few papers addressing the potential role of capacity payments as a way to solve the underinvestment problem from a formal perspective, namely Brown (2014), Elberg and Kranz (2013) and Schäfer and Schulten (2014) (for more information about the literature, see the previous chapter).

²This mechanism is formally called the “Operating Reserve Demand Curve”. For more information, see Hogan (2013).
³See, for instance, Ausubel and Cramton (2010), Cramton and Stoft (2005) and Joskow (2008b).
In this chapter we provide a formal model that can be used to study capacity payments and their impact on consumers’ welfare. We propose a two-stage environment in which two types of electricity generators –base load and peak load– invest in capacity in the first stage (before demand is realized) and then compete in a wholesale market to produce and sell electricity to consumers.\footnote{In this chapter, we do not discuss the role of distributed generation. We recognize that this is an important caveat.}

Applying backward induction, we first deal with the second stage. As in Fabra et al. (2006), we propose a uniform-price auction as the mechanism to allocate electricity in the wholesale market.\footnote{Uniform-price auctions are consistent with many real-world market. In fact, many authors use this kind of auctions in a similar context to ours. See, for instance, ?, Von der Fehr and Harbord (1998), Fabra et al. (2006) and Brown (2014). Further motivation is provided in the previous chapter.} In the first stage, electricity generators solve for the equilibrium investment in capacity by considering the second-stage expected profits. Then, using the equilibrium that we find, we analyze consumers’ welfare, measured in terms of expected consumer surplus. We also analyze the consequences in terms of the expected amount paid by consumers, price volatility and system reliability.

We investigate two benchmark cases. First, we characterize the competitive industry outcome.\footnote{We employ usual “competitive industry” assumptions –see, for instance, Tremblay and Tremblay (2012), pp. 123-143.} In this setup, there is a very large number of generators that serve the market in the second stage. As is standard, at the time of investing in capacity (the first stage) there is free-entry and the generators expect (long-run) zero-profit in equilibrium. Although the case of perfect competition lacks of realism, this analysis provides a paradigmatic benchmark. Second, to allow for certain degree of market power in the electricity market, we assume two dominant firms facing a competitive fringe. As Bonacina and Gulli’ (2007) point out in a similar context, the dominant-fringe model is suited to represent the reality of several electricity markets (including the Texas ERCOT market itself).\footnote{In many liberalized and restructured markets, the former electricity monopolist usually retains a greater share of the market, and the new-born entrant firms are usual smaller players. In fact, this framework is becoming increasinly popular in the electricity markets literature. See Ito and Reguant}
of two main players (TXU and Reliant) and many small firms, local utilities and cooperatives that form the competitive fringe.

We obtain the following results. First, in the context of an “energy-only” market, we show that price caps are not as beneficial as previously thought. In particular, we find that, in a competitive setting, an increase in the cap induces a substantial increase in consumer surplus.\(^8\) Second, we show that the introduction of a capacity market is able to reduce the probability of a blackout (it increases system reliability) and reduce price volatility. It does so, however, at the cost of increasing the expected amount paid by consumers in the electricity market. This is due to the fact that the capacity payments are effectively passed-through to consumers via higher power prices. Such an increase in the average amount paid by consumers leads to a reduction in consumer surplus in both a competitive market scenario and in the presence of some degree of market power.

The contributions of this chapter are twofold. First, on the theoretical side, we extend the 2-firm framework in Fabra et al. (2006) to the general (heterogeneous) \(n\)-firm case.\(^9\) Such an extension, which, to our knowledge, is absent in the literature, is well-suited to study both a competitive market and a market with the presence of dominant firms. Up to now, most of the literature has focused on the two-firm case. While this was a reasonable assumption 10 or 15 years ago, a duopoly does not seem to capture the reality in many of the current restructured and liberalized electricity markets. On the policy side, our aim is to provide a contribution by studying the impact of capacity compensation mechanisms on consumers’ welfare.

The rest of the chapter is organized as follows. Section 3.2 provides the theoretical

\(^8\)In the “dominant-fringe” model, however, we find that an increase in the price cap reduces consumer surplus.

\(^9\)We are aware that Fabra et al. (2006) provide an extension of their duopoly model for the oligopoly case. However, we want to go beyond that, by providing a complete game-theoretic framework, including an allocation rule, a profit rule and the Nash Equilibrium (NE) strategies.
framework that we use throughout the chapter. In section 3.3 we characterize the equilibrium strategies in the second stage—the electricity market. Section 3.4 builds on section 3.3 results to determine the equilibrium investment in capacity. We obtain an equilibrium both with and without capacity compensation payments. In section 3.5 we present the results for the “dominant-fringe” model. Section 3.6 describes the welfare measures that we use in section 3.7 to evaluate the ERCOT market in different counterfactual scenarios. Finally, section 3.8 concludes. All proofs are in Appendix B.

3.2 Theoretical framework

There are $n_b$ risk neutral identical base load electricity generators (b) and $n_p$ risk neutral identical peak load electricity generators (p). Both types of suppliers invest in capacity, and produce electricity up to the amount of their capacity. They compete to sell electricity to consumers in a wholesale market.

We assume that consumers’ demand takes the form of a rotated “L”. In particular, each consumer has a reservation price which is technically called Value of Lost Load (VOLL). It reflects consumers’ opportunity cost of having an interruption of electricity supply during a “blackout”.

For the sake of simplicity, we assume that there is unit one continuum of consumers, whose reservation price (VOLL) is denoted as $p^H > 0$. However, the price $p^H$ is usually not achieved. This is true because in most (if not all) current liberalized markets, prices are capped. In other words, there is a maximum price above which the regulator does not allow power generators to sell electricity in the wholesale market. In our simple model, we say that the price cap, denoted as $p^{\text{cap}}$,

---

10 We focus on the case of two groups for two main reasons. First, because our aim is to study the Texas ERCOT case, which has been dominated by two generation technologies (coal and natural gas). Second, for the sake of expositional clarity. However, the results that we present here apply also to the case in which there are more than two types of generators.

11 This is not a system-wide loss of supply but rather a curtailment of supply, for example, to customers with interruptible supply contracts. See Cramton et al. (2013), Spees et al. (2013).
is binding if \( p^\text{cap} \in [0, p^H) \), and is not binding –it has not effect on the consumers– if \( p^\text{cap} \in [p^H, +\infty) \). Let us denote \( \bar{p} \equiv \min\{p^H, p^\text{cap}\} \).

Consumers’ aggregate demand, \( \theta \), is stochastic. In particular, we assume that \( \theta \) is a non-negative random variable distributed according to some known cumulative distribution function \( F(\theta) \), defined for the support \([\bar{\theta}, \bar{\theta}]\). Without loss of generality, we normalize this support to \([0,1]\). Furthermore, we assume that \( F(\cdot) \) is strictly increasing, continuous and \( F^{-1}(\cdot) \) exists.\(^{12}\)

The timing in this market is as follows:

1. Each generator \( j \) decides how much to invest in capacity, i.e. \( k^j_i \geq 0, i \in \{b, p\} \).\(^{13}\)

2. Demand is realized, \( \theta \in [\bar{\theta}, \bar{\theta}] \). Then, each supplier simultaneously submits a bid specifying the minimum price at which it is willing to produce and sell electricity in a centralized wholesale market.\(^{14}\)

When investing in capacity, generators incur a cost. In particular we assume that electricity generators face constant linear costs \( c_k_i > 0 \) of building \( k_i \geq 0 \) units of capacity, \( i \in \{b, p\} \). We assume that \( c_{k_b} > c_{k_p} \).\(^{15}\)

Generator \( j \) produces electricity out of capacity according to the following production function:

\[
0 \leq q^j_i \leq k^j_i \quad i \in \{b, p\}
\] (3.1)

i.e. production above capacity is impossible. Both \( k^j_i \) and (obviously) \( q^j_i \) are assumed to be perfectly divisible.

\(^{12}\)This kind of inelastic random demand is widely used in the literature examining wholesale electricity markets. See, among others, Bonacina and Gulli’ (2007), Crampes and Creti (2005) and Fabra et al. (2006).

\(^{13}\)More precisely, there is a previous step in which a large (infinite) number of firms decide whether to enter or not the market. The analysis of this stage can be included in this analysis, but little is added by so doing. Hence, we omit this step and we assume that \( n_i \) is the free-entry equilibrium number of firms, for \( i \in \{b, p\} \).

\(^{14}\)This assumption is based on the fact that in wholesale (day-ahead) markets, the following day system load is almost perfectly assessed, with some minimal differences that are matched usually through balancing and/or ancillary services markets. In contrast, at the time of investing in generation capacity, which usually takes several months, the demand is totally unpredictable.

Electricity is traded in the wholesale market at price $p^s > 0$. Base load generators incur a variable cost of $c_b > 0$ for producing electricity out of capacity while for peak load generators the variable cost is $c_p > c_b$. For both types of generators, production above capacity is impossible (infinitely costly). The assumption that $c_b < c_p$ captures the usual so-called merit order observed in electricity markets.\(^\text{16}\)

Let us denote $k_i = \sum_{j=1}^{n_i} k_i^j$ the total capacity installed by generators of type $i \in \{b, p\}$. In addition, we denote the total installed capacity in the market as $K$, where $K = k_b + k_p$. Thus, once the demand $\theta$ is realized, we face three possible scenarios:\(^\text{17}\)

1. $k_b \geq \theta$. That is, base load capacity is greater than aggregate demand. Thus, base load generators have enough capacity to serve the whole demand.

2. $k_b < \theta < K$. That is, base load capacity is not enough to serve the demand, but the whole demand can be served by both type of generators.

3. $K \leq \theta$. That is, there is not enough capacity to serve the whole demand.

Next, as Creti and Fabra (2007) do, we proceed by backward induction. Thus, in a similar way as we do in the previous chapter, we first characterize the Nash equilibrium bidding strategies for both base load and peak load generators in the wholesale market given the capacity choices. Then, given the equilibrium bids and profits, we obtain the equilibrium capacities.

### 3.3 Equilibrium in the wholesale electricity market

Given the capacities that each generator chooses in the first period, $k_i^j$ for $j \in \{1, \ldots, n_i\}$ and $i \in \{b, p\}$, these suppliers compete in a centralized uniform-price

---

\(^\text{16}\)To avoid uninteresting cases, we further assume that $c_i < \bar{p}$ for $i \in \{b, p\}$.  

\(^\text{17}\)We restrict ourselves to the cases in which $k_i^j \geq 0$ for some $i \in \{b, p\}$ and $K \leq 1$.  

auction\textsuperscript{18} to supply electricity to consumers. In particular, once they observe the realization of demand, electricity generators simultaneously submit an offer price (bid), $p^i_j > 0$. Through an offer price, a generator specifies the minimum price at which it is willing to produce and sell electricity. For the sake of simplicity we assume that each generator submits a single price offer for its entire capacity.\textsuperscript{19} Thus, a strategy for generator $(i, j)$ is simply given by a rule $p^i_j : [0, 1] \rightarrow \mathbb{R}_{++}$. 

An independent agent clears the auction as follows. The lowest-bidding generator is dispatched (and thus serves the market) first. Higher-bidding generators are dispatched subsequently (according to their bids, in an increasing way) to serve the residual demand if the combined capacity from the lower-bidding generators is not enough to serve the realized aggregate demand. In case of a tie, the rationing rule is just pro-rata among marginal bidders whose bids are equal (“pro-rata on the margin”).\textsuperscript{20} Analytically, the quantity allocated to generator $(i, j)$ is:

$$q^i_j(p^*; p^i_j(\cdot), \theta) = \begin{cases} k^i_j & \text{if } p^i_j < p^* \\ \mathbbm{1}_{|S| = 1} \left\{ \min\{k^i_j, \theta - \sum_{l \in L} k^l_j\} \right\} + \mathbbm{1}_{|S| \geq 2} \left\{ R(p^*; p^i_j(\cdot), \theta) \right\} & \text{if } p^i_j = p^* \\ 0 & \text{if } p^i_j > p^* \end{cases}$$

where $S$ is the set of firms posting the market-clearing price, $p^*$, including firm $(i, j)$, $L$ is the set of firms posting a price (strictly) below $p^*$, and $R(p^*; p^i_j(\cdot), \theta)$ is the usual “pro-rata-on-the-margin” rule –see Kremer and Nyborg (2004), among many others.

The market-clearing price, $p^*$, is the minimum price at which aggregate demand equals generators’ supplied capacity. If aggregate demand is greater than firms’ ag-

\textsuperscript{18}See Newberry (2002) and Von der Fehr and Harbord (1998).

\textsuperscript{19}Although in reality each generator can submit a bid function specifying multiple prices and quantities, as Fabra et al. (2006) show, this assumption simplifies the analysis and it is largely inessential. This is also equivalent to allowing generators to submit a price function contingent on realized demand.

\textsuperscript{20}In the context of uniform-price auctions, Kremer and Nyborg (2004) show that tie-breaking rules matters for the market equilibrium. However, we show that our results are independent of the tie-breaking rule.
aggregate capacity, the market is cleared at $\bar{p}$. Formally,

$$p^*(\theta) = \begin{cases} 
\min\{p \leq \bar{p} | \sum_{j=1}^{n_b+n_p} q_i^j(p; p_i'(.), \theta) = \theta\} & \text{if } \{p \leq \bar{p} | \sum_{j=1}^{n_b+n_p} q_i^j(p; p_i'(.), \theta) = \theta\} \neq \emptyset \\
\bar{p} & \text{if } \{p \leq \bar{p} | \sum_{j=1}^{n_b+n_p} q_i^j(p; p_i'(.), \theta) = \theta\} = \emptyset 
\end{cases} \quad (3.3)$$

where $\sum_{j=1}^{n_b+n_p} q_i^j(p; p_i'(.), \theta)$ is aggregate production at price $p$, given firms’ strategies and realized aggregate demand.

Since the quantity allocation mechanism is a uniform-price auction, for each unit of electricity produced and sold, generator $(i, j)$ receives $p_s$. Thus, the payoff (profit) for supplier $(i, j)$, denoted as $\pi^j_i$, is given by:

$$\pi^j_i(p^s; p_i^j(.), \theta) = (p^s - c_i) \cdot q_i^j(p^s; p_i^j(.), \theta) - c_{k_i} k_i^j \quad (3.4)$$

Given this environment, the following proposition characterizes the Nash equilibrium strategies of the auction game for different scenarios and types of generators.

**Proposition 6. Weakly-dominant equilibrium strategies.** Given electricity generators’ capacities$^{21}$, $k_i^j$ for $j \in \{1, \ldots, n_i\}$, $i \in \{b, p\}$, in every pure-strategy equilibrium:

1. if $\theta \leq \left(\frac{n_b-1}{n_b}\right) k_b$, $p_b^j \leq c_b$, with $p_b^{j'} = c_b$ for some $j'$, which clears the market, and $p_p^j \geq c_b \forall j$.

2. if $\left(\frac{n_b-1}{n_b}\right) k_b < \theta < k_b$, $\exists! j'$ such that $p_b^j = c_p$, $p_b^{j'} = \epsilon > 0 \forall j \setminus \{j'\}$, and $p_p^j \geq c_p \forall j$.

3. if $k_b < \theta \leq \left(\frac{n_b+n_p-1}{n_b+n_p}\right) K$, $p_b^j < c_b$ and $p_p^j > p_b^j \forall j$, with $p_p^{j'} = c_p$ for some $j'$, which clears the market.

4. if $\left(\frac{n_b+n_p-1}{n_b+n_p}\right) K \leq \theta < K$, $\exists! j'$ such that $p_i^{j'} = \bar{p}$ and $p_i^j = \epsilon_i > 0 \forall j \setminus \{j'\} \forall i$.

$^{21}$Notice that, according to the symmetric nature of the generators, capacity levels are assumed to be symmetric for every type $i$ generator.
5. if $K \leq \theta$, $p_i^j \leq \bar{p} \forall i, j$, with $p_p^j = \bar{p}$ for some $j'$, which clears the market.

Proposition 6 characterizes the (weakly-dominant) equilibrium strategies for all the possible realizations of $\theta$ relative to the capacity levels.\footnote{In this Proposition we are missing the knife-edge case in which $\theta = k_b$. In such case, it is trivial that in every pure-strategy equilibrium $p_b^j \leq c_p \forall j$, with $p_b^{j'} = c_p$ for some $j'$ and $p_p^j > c_p \forall j$.} We comment on a few interesting facts in this Proposition. Cases #1 and #3 are not surprising. In these cases, there is excess capacity for more than one base load generator (case #1), and for more than one generator, of any type (case #3). Thus, generators compete à la Bertrand, yielding the well-known result that the market clearing price is equal to the marginal cost. Case #5 is also not surprising. In this case, capacity is scarce and consumers are “ripped off”, i.e. generators are able to impose the maximum price.

Scenarios #2 and #4 are special though. In case #2, the only possible equilibrium is achieved when all the base load generators but not one ($j'$) bid a low amount $\epsilon^j > 0$, and a base load generator $j'$ bids exactly at the peak load marginal cost, while peak load generators bid above their marginal costs. By bidding at $\epsilon^j > 0$, small enough, all other base load generators prevent $j'$ to deviating from serving less than the residual demand to serve a greater portion of the demand at a bid lower or equal to $\epsilon^j$. If so, $c_p$ will no longer be the market clearing price, but it will be $\epsilon^j$, making this deviation unprofitable for $j'$. The same logic applies to case #4.\footnote{Considering that we are looking for weakly dominated strategies, for a peak load generator there is no strictly profitable deviation.}

### 3.3.1 Perfect competition in the wholesale electricity market

Consistent with the perfectly competitive benchmark case, in this subsection we assume that the number of electricity generators of both types, $n_i$, is large enough. More precisely, we consider the limiting case in which there are infinitely many non-atomic base load and peak load generators of infinitesimal size. We index them by $j$, where $0 \leq j < 1$. Moreover, in this scenario, type-$i$ generators’ aggregate capacity,
$k_i$, is given by $k_i = \int_0^1 k_i^j dj$ and type-$i$ generators’ aggregate profit, $\pi_i$, is given by $\pi_i = \int_0^1 \pi_i^j dj$.

This assumption not only allows us to study the benchmark case of perfect competition. It also avoids the “knife-edge” cases included in Proposition 6 (case #2 and case #4) in which $\theta$ is less than $k_b$ but close enough to $k_b$ (case #2) and in which $\theta$ is less than $K$ but close enough to $K$ (case #4) such that the “last-dispatched” generator exercises some market power by clearing the market above its marginal cost. Thus, in a perfect competition context, $n_i \to \infty$ guarantees the existence of a Bertrand-like equilibrium, which is something one should expect in a perfectly-competitive environment. This unsurprising result is embedded in the following proposition, in which we solve for the price, production and (ex-post) wholesale profit (payoffs) in equilibrium, in a similar ways as we did in the first proposition of the previous chapter.

**Proposition 7.** Given electricity generators’ capacities, $k_i$ for $i \in \{b, p\}$, in a perfectly competitive market:

1. If $k_b \geq \theta$ the equilibrium spot market price is $p^s = c_b$ and both generators earn zero profit ($\pi_i = 0$).

2. If $k_b < \theta \leq K$ the equilibrium spot market price is $p^s = c_p$. Base load generators’ profits are positive ($\pi_b > 0$) and peak load generators earn zero profit ($\pi_p = 0$).

3. If $K < \theta$ the equilibrium spot market price is $p^s = \bar{p}$ and both generators earn positive profits ($\pi_i > 0$).

The intuition behind this result is the following. If base load capacity is greater than aggregate demand ($k_b \geq \theta$), base load generators have enough capacity to serve the whole demand, and competition drives the price to their marginal cost, $c_b$. Hence $p^s = c_b$. If base load capacity is not enough to cover the demand ($k_b < \theta \leq k_b + k_p$) then there is room for peak load generators to produce and sell electricity. Again,
competition drives the price to their marginal cost, \( c_p \), and both types of generators receive the same \( p^* \), which is given by \( p^* = c_p \). In the third case, capacity is scarce. Hence, (excess) demand drives the price up either to \( p^H \), which is equal to the price consumers are willing to pay to avoid an interruption of power supply (VOLL), or to the price cap \( (p^\text{cap}) \), if binding. In both cases, the part of the demand that is not served suffers a power outage.

Although it is useful and theoretically enriching to analyze a paradigmatic benchmark case, we cannot ignore that assuming that firms behave competitively might be unrealistic. Thus, in section 3.5 we abandon this framework and characterize a new equilibrium when we assume instead that there exists a certain degree of market power. In particular, we assume that there are two dominant generators (a base and a peak) that face a competitive fringe. This model is closer to actual markets such as, for instance, the ERCOT market in Texas – see Bonacina and Gulli’ (2007).

### 3.4 Equilibrium capacities

Next, let us examine the problem that both types of electricity generators solve in the first period. According to the previously described timing, supplier \( j \) decides in this period how much to invest in capacity, \( k^j_i \) for \( i \in \{b, p\} \). This decision is taken before knowing the realization of the demand. However, the capacity levels are chosen anticipating the equilibrium profits in the centralized electricity market.

As seems reasonable, we restrict the attention to the cases in which \( k^j_i \in [0, 1] \), for all \( (i, j) \). On the one hand, at least in a static model, it does not make sense to talk about negative investment (divestment) since generators enter period one with no installed capacity at all. On the other hand, it also seems unreasonable in our basic framework for the generators to choose a level of investment greater than 1, since investment is costly for them and part of the capacity would then always be
In this section we study both the equilibrium capacities without capacity payments and the equilibrium capacities in the presence of capacity payments. According to Creti and Fabra (2007), there are two main types of capacity payment systems: price-based systems and quantity-based systems. In a price-based system, generators decide how much capacity they will make available given a price that is set by the market regulator. In a quantity-based system, the regulator fixes a certain amount of capacity—which is typically equal to the forecasted peak demand plus a reserve margin—and generators compete through some form of market mechanism to provide this level of capacity in return for some compensation.\footnote{A broader classification is in Spees et al. (2013). They include the so-called price-based system in the group of “Administrative mechanisms”, while the quantity-based systems are included in the group of “Market-based mechanisms”. Furthermore, they distinguish between the “LSE RA requirement systems”, in which the capacity allocations are achieved through bilateral agreements (e.g. CAISO); and the (proper) “Capacity Markets”, in which centralized auctions are used (e.g. PJM, NE-ISO, NYISO).} Price-based systems have barely been analyzed in the past—see, for instance, Newbery (1995) and Wolak and Patrick (2001)—and have been excluded from the current debate. Thus we focus our analysis on the quantity-based systems, which are actually the ones that the aforementioned countries and regions are considering. In these systems, the regulator sets an administrative reliability requirement (also called resource adequacy requirement) that the retailers must guarantee to achieve. This administrative reliability requirement amounts to a level of capacity that the agents in the market must guarantee and such that the whole demand is (almost surely)\footnote{A common rule of thumb is to calculate the necessary capacity to achieve a probability of blackout no larger than once in every 10 years. See Spees et al. (2013), footnote 7.} covered at a certain point in the future.\footnote{If this capacity level is not achieved, the regulator has the right to impose fines on the defaulting firms. See Crampes and Creti (2005).}

In our setup, this mechanism is modeled as follows. The regulator fixes a (very low) probability level of blackout, denoted as $\eta^T \geq 0$ or, equivalently, a (strictly) positive target capacity investment $K^T$, such that $\eta^T \equiv Pr(K^T \leq \theta)$. In particular, $K^T$ must
be equal to the sum of the total installed capacity that both types of generators make in equilibrium. In other words, \( K^T = k^{*,m}_b + k^{*,m}_p \), where \( k^{*,m}_i \) is the equilibrium aggregate investment in capacity made by suppliers of type \( i \in \{b,p\} \) provided that there is a compensation payment, \( m \). To achieve this target, the electricity generators that invest in capacity receive a compensation \( m \) per-unit of capacity built. This compensation is paid by consumers.\textsuperscript{27}

### 3.4.1 Equilibrium capacities in a perfectly competitive market

a) Scenario #1: equilibrium in the absence of capacity payments

To continue with the analysis of the benchmark case, free-entry is assumed in the first stage. More precisely, we assume that the number of firms that enter period 1 is the free-entry equilibrium number of firms. That is, \( n_i \) is such that \( E\pi_i \geq 0 \) if the number of firms is \( n_i \), and \( E\pi_i < 0 \) if the number of firms is \( n_i + r \), for \( r = 1, 2, \cdots \), where \( E\pi_i \) is the (expected) profit for firm \((i,j)\) at period 1. As Mankiw and Whinston (1986) do, and considering that the number of potential entrants is large and their size is infinitesimally small (perfect competitors), we ignore the integer constraint on the number of firms. Thus, we assume that the free-entry equilibrium number of suppliers \( (n_i) \) exactly satisfies the zero-profit condition, i.e. \( E\pi_i = 0 \).

Therefore, considering that the number of firms is sufficiently large (perfect competition), the following proposition (which is similar to the second proposition in the previous chapter) captures the long-run equilibrium capacities in the absence of a capacity compensation mechanism.

**Proposition 8.** *Equilibrium capacities in a competitive market in the absence of capacity payments.* Denote \( a_i \equiv \bar{p} - c_i, \ i \in \{b,p\} \). In a perfectly

\textsuperscript{27}Typically retailers pay such compensations. However, as explained by several authors – Arizu et al. (2004), Hunt (2012) and Thomas et al. (2014) among others– these compensations are effectively passed-through to final consumers (via higher electricity prices).
competitive market at an interior solution the unique aggregate base load capacity is 

\[ k^*_b = F^{-1} \left( 1 - \frac{c_{kb} - c_{kp}}{c_p - c_b} \right) \]

and the unique aggregate peak load capacity is 

\[ k^*_p = F^{-1} \left( 1 - \frac{c_{kp}}{c_p} \right) - F^{-1} \left( 1 - \frac{c_{kb} - c_{kp}}{c_p - c_b} \right). \]

We comment on a couple of ancillary results from Proposition 8. First, even if both types of suppliers invest a positive amount in capacity, aggregate installed capacity in the market is less than 1 (the maximum aggregate demand). Second, it also follows immediately from Proposition 8 that the introduction of a binding price cap provides a disincentive to invest in generation capacity. We formally capture these ideas in the following corollaries.

**Corollary 3.** In equilibrium, aggregate installed capacity is strictly less than one. That is, 

\[ K^* \equiv k^*_b + k^*_p < 1. \]

Corollary 3 implies that in the absence capacity payments, a competitive market does not achieve the investment level necessary to avoid blackouts if the demand is sufficiently (and extraordinarily) high.

**Corollary 4.** Let us denote \( k^*_i \) type-\( i \) generators’ equilibrium aggregate capacity and \( K^* \) equilibrium aggregate capacity in a market with a non-binding price cap, and \( k^*_{i,\text{cap}} \) generators’ equilibrium aggregate capacity and \( K^*_{\text{cap}} \) equilibrium aggregate capacity in a market with a binding price cap. Then 

\[ k^*_{i,\text{cap}} \leq k^*_i, \ i \in \{b,p\} \text{ and } K^*_{\text{cap}} < K^*. \]

According to Corollary 4, the introduction of a binding price cap in the market, i.e. \( p_{\text{cap}} \in (c_p, p^H) \), reduces the equilibrium capacity investment relative to a market with no price cap. This is consistent with the result found (among others) by Joskow and Tirole (2007) and with economic intuition. Since a price cap reduces generators’ rents—which are obtained during scarcity events– it makes them less willing to incur

\[ \text{---} \]

\[ 28 \text{ Notice that in all cases the equilibrium capacity investment of generator } i \text{ depends on the costs of generator } i', i, i' \in \{b,p\}, i \neq i'. \text{ However, as extensively discussed in Crampes and Creti (2005), Hortacsu and Puller (2008) and Inal (2011) (among others), firms usually have access to rivals' information about costs since they all use essentially the same technologies.} \]
investment costs, resulting in less installed capacity. The same idea is in Hogan (2005), who point out that price caps indeed lead to underinvestment in capacity (the “missing money” problem).

b) Scenario #2: equilibrium in the presence of capacity payments

As discussed above, in the presence of a capacity adequacy requirement and capacity payments, the regulator fixes a capacity target $K^T$ that must be achieved. To achieve this target, the electricity generators that invest in capacity receive a compensation $m$ per unit of capacity built. With this new setting in mind, the next proposition solves for the per-unit compensation payment that generators are willing to accept to achieve an equilibrium in which the investment in capacity is given by $k_i^{*,m}$, $i \in \{b, p\}$, such that $K^T = k_b^{*,m} + k_p^{*,m}$.

Proposition 9. Equilibrium capacities in a competitive market in the presence of capacity payments. Denote $a_i \equiv \bar{p} - c_i$, $i \in \{b, p\}$. Given $K^T \in (0, 1]$, in a perfectly competitive market at an interior solution the unique aggregate base load capacity is $k_b^{*,m} = F^{-1} \left( 1 - \frac{c_b - c_{kp}}{c_p - c_b} \right)$, the unique aggregate peak load capacity is $k_p^{*,m} = K^T - F^{-1} \left( 1 - \frac{c_b - c_{kp}}{c_p - c_b} \right)$, and the unique per-unit equilibrium capacity compensation payment is $m = c_{kp} + a_p [K^T - 1]$.

From this Proposition (which is similar to the third proposition in the previous chapter) it follows that, in the presence of capacity payments, the regulator is able to solve the underinvestment problem by setting a sufficiently high capacity compensation payment.

Corollary 5. There exists an $m$ such that $K^T > K^*$. 

In the numerical analysis section we discuss that although a capacity compensation mechanism is able to solve the underinvestment problem, it may raise on average the
price paid by electricity consumers. This is due to the fact that, as discussed above, the capacity compensation is effectively passed-through to final consumers.

Through the rest of the chapter we use the previous results (Propositions 7, 8 and 9) to analyze the welfare implications for consumers in a perfectly competitive electricity market. Again, in section 3.5, we partially abandon the competitive framework and we assume instead that two dominant generators (a base load and a peak load) choose capacities to maximize their expected profits once they observe aggregate capacity choices of a set of competitive firms.

### 3.5 Market power in electricity markets: dominant firm-competitive fringe model

In this subsection, we propose a dominant firm-competitive fringe model, following closely the formulation proposed first by Shitovitz (1973) and then extended by Okuno et al. (1980). In the context of the electricity sector, this market structure is also employed by Elberg and Kranz (2013), Ito and Reguant (2014), Tanaka and Chen (2013) and Vespucci et al. (2013), among others. As Bonacina and Gulli’ (2007) state, the dominant-fringe framework represents the reality of several electricity markets, in which large and well-recognized power firms –typically former public monopolists– still control a greater share of the market relative to the new firms that entered the market after the liberalization phase.

We characterize the Subgame Perfect Nash Equilibrium (SPNE) strategies in the two-stage game for two dominant suppliers, a peak load generator and a base load generator, facing a competitive fringe.\textsuperscript{29} Regarding the competitive fringe, we consider

\textsuperscript{29}In general, a strategy profile is a SPNE if it represents a Nash Equilibrium (NE) of every subgame of the game. In our particular context, by using SPNE as our solution concept, we guarantee that the strategies in both the capacity stage and the auction stage constitute a NE in each of these stages.
the limiting situation in which there are infinitely many base load generators (b) and infinitely many peak load generators (p). These generators invest in capacity in the first stage and compete to procure electricity to consumers in a wholesale market. Whenever the competitive fringe serves the market (and provided that there exist some excess capacity), competition between firms drives the market price down to the marginal cost.

On the other hand, the dominant generators act strategically. In particular, in the first stage, they choose their investment capacity levels \( (k^d_i, \text{ for } i \in \{b, p\}) \) once they observe the “fringe-generators” capacities. In other words, the dominant generators act as Stackelberg leaders in the first stage. The dominant firms choose the capacities by maximizing their expected stage-2 profits, anticipating how the competitive fringe will react.

Let \( \int_0^1 k^f_i dj = k^f_i \) denote non-dominant generators’ aggregate capacity and \( k_i = k^f_i + k^d_i \) type-\( i \) aggregate capacity. Furthermore, let us denote aggregate capacity in the market as \( K \), where \( K = k_b + k_p \). Finally, let us denote type-\( i \) fringe generators’ aggregate profit as \( \pi^f_i \), where \( \pi^f_i = \int_0^1 \pi^f_i dj \), and type-\( i \) dominant generator’s profit as \( \pi^d_i \). Assumptions regarding demand, costs, and production remain unchanged.

Again, we proceed by backward induction. We first solve the wholesale market game, and then we find the equilibrium capacities.

---

\(^{30}\)In other words, there is a continuum of small base and peak load generators. These generators represent in actual markets small firms, cooperatives and municipal utilities.

\(^{31}\)This assumption does not literally imply that dominant generators actually build whatever capacity is optimal such that they maximize their profit given fringe generators’ capacities. Our goal is to capture the idea that dominant generators’ capacities may play a strategic role, for instance, through withholding. This idea is in Crampes and Creti (2005) and LeCoq (2002). In fact, the former author analyzes the strategic use of generation capacity in the electricity market using a two-firm model.

\(^{32}\)A reviewer suggested the inclusion of a cost advantage for the dominant firm relatively to the competitive fringe. Although this assumption seems reasonable (it captures the existence of economies of scale), the equilibrium obtained in the second stage is observationally equivalent to the equilibrium presented in this section in the absence of a dominant firm’s cost advantage.
3.5.1 Equilibrium in the wholesale electricity market

Given generators’ capacities, $k_i^f$ and $k_i^d$, $i \in \{b, p\}$, generators submit an offer price (bid) once they observe the realization of the demand, $\theta$. In this section we focus on the behavior of the dominant generators (the strategic players). A strategy for a dominant generator $(i, d)$ is simply given by a mapping $p_d^i : [0, 1] \rightarrow \mathbb{R}^+$. Again, the allocation rule is based on a uniform-price auction and the market-clearing price and firms’ profits are as characterized in section 3.3.

Having this setup in mind, the following proposition characterizes the equilibrium strategies for the dominant firms facing a very large number (infinite) of small competitive generators that form a competitive fringe.

Proposition 10. **Dominant firms’ equilibrium strategies.** Given electricity generators’ capacities, $k_i^f$ and $k_i^d$, $i \in \{b, p\}$, in every pure-strategy equilibrium:

1. If $k_b^f \geq \theta$, $p_b^d \geq c_b$ and $p_p^d > c_b$.  
2. If $k_b^f < \theta \leq k_b^f$, $p_b^d = c_p$ and $p_p^d > c_p$.  
3. If $k_b^f < \theta \leq k_b^f + k_p^f$, $p_b^d \leq c_p$ and $p_p^d \geq c_p$.  
4. If $k_b^f + k_p^f \leq \theta$, $p_i^d \leq \bar{p}$.

The intuition in this equilibrium is the following. If the competitive fringe serves the market, competition drives the price to the marginal cost. However, if a dominant firm becomes the “marginal producer”, it is able to exercise market power and raise the price up to either the marginal cost of peak load generators (case #2) or to $\bar{p}$ (case #4).

---

33 As discussed above, whenever there is “fringe strictly positive production” and “fringe excess capacity”, competition will drive the price down to the marginal cost.
34 For the sake of completeness, notice that $p_b^d \in \mathbb{R}^+$ if $k_b^d < \theta$.
35 In this case, it is assumed that peak load fringe generators do not serve the market. Otherwise, there is no equilibrium.
36 For the sake of completeness, notice that $p_p^d \in \mathbb{R}^+$ if $k_b^f + k_b^d + k_p^d < \theta$. 
Next, using the equilibrium strategies in Proposition 10, the following corollary captures the market-clearing price and profits in equilibrium.

**Corollary 6.** Given electricity generators’ capacities, \( k_i^f \) and \( k_i^d \), \( i \in \{b,p\} \):

1. If \( k_b^f \geq \theta \) the equilibrium market-clearing price is \( p^s = c_b \) and all generators earn zero profit (\( \pi_i^d = \pi_i^f = 0 \ \forall i \)).

2. If \( k_b^f < \theta < k_b^f + k_p^f \) the equilibrium market-clearing price is \( p^s = c_p \), base load generators’ profits are positive (\( \pi_b^f, \pi_b^d > 0 \)) and peak generators earn zero profit (\( \pi_p^f = \pi_p^d = 0 \)).

3. If \( k_b^f + k_p^f \leq \theta \) the equilibrium market-clearing price is \( p^s = \bar{p} \) and all generators earn positive profit (\( \pi_i^d, \pi_i^f > 0 \ \forall i \)).

This result, which will be a key one when dealing with the first stage, is in line with the explanation given above for Proposition 10.

### 3.5.2 Equilibrium capacities

Next, we examine the problem that both types of electricity generators solve in the first period. According to the market timing, in this period each supplier chooses how much capacity to invest in. This decision is taken before knowing the realization of the aggregate demand. However, the capacity levels are chosen considering the expected equilibrium profits in the wholesale spot electricity market (Corollary 6).

Again, we focus on the behavior of the dominant generators (the strategic players).\(^{37}\) A strategy for a dominant generator \((i, d)\) is \( k_i^d(k_b^f, k_p^f; k_{-i}^d) \in [0, 1] \). In other words, the dominant generators choose the level of investment once they observe the capacity investment made by the competitive fringe.

\(^{37}\)Regarding the competitive fringe, we maintain the same “competitive market” assumptions employed in Section 3.3.
As in Section 3.4, in this section we study both the equilibrium capacities in the absence of capacity payments and the equilibrium capacities in the presence of capacity payments. In the latter case, as we did in section 3.4, we consider a quantity-based system mechanism. That is, a system in which the regulator fixes a (strictly) positive capacity target $K^T \leq 1$ (the resource adequacy requirement) and the electricity generators that invest in capacity receive a compensation $m$ per unit of capacity built.

a) Scenario #1: equilibrium in the absence of capacity payments

First, we consider the case in which there are no capacity compensation payments in the market. In this case, we first find the equilibrium investment in capacity made by the competitive fringe generators. Recall that for these generators, we have imposed the usual “competitive firms” assumptions (see section 3.3). Having these assumptions in mind, the equilibrium aggregate capacity investment of the fringe firms is included in the following proposition.

**Proposition 11.** *Equilibrium capacities chosen by the competitive fringe in the absence of capacity payments.* Denote $a_i \equiv \bar{p} - c_i$, $i \in \{b,p\}$. At an interior solution, the base load fringe’s equilibrium aggregate capacity is $k_{b,f}^* = F^{-1}\left(1 - \frac{c_{lb} - c_{lb}}{c_{lb} - c_{l_b}}\right)$ and the peak load fringe’s equilibrium aggregate capacity is $k_{p,f}^* = F^{-1}\left(1 - \frac{c_{lp}}{a_p}\right) - F^{-1}\left(1 - \frac{c_{lp} - c_{lp}}{c_{lp} - c_{l_b}}\right) - k_{b}^d$.

The competitive fringe equilibrium aggregate capacity is anticipated by the dominant firms. A dominant generator $(i,d)$ chooses the period-2 profit-maximizing level of investment in capacity anticipating the fringes’ equilibrium aggregate capacity investment and given $(-i,d)$ dominant generator’s capacity investment. That is, generator $(i,d)$’s optimal capacity investment solves:

$$k_{i,d}^* = \arg\max_{k_{i,d} \in [0,1]} E\pi_i^d(p^*, p_{i,(\cdot)}^d, \theta)$$ (3.5)
where $E$ is the expectation operator at period 1 and $\pi^d_i(p^*_s; p^I_i(\cdot), \theta)$ denotes generator $(i, d)$’s profit.\footnote{For the optimization routine, we employ a constrained optimization routine using R.}

**Proposition 12.** Dominant generators’ equilibrium capacities in the absence of capacity payments. $k^*_i \in [0, 1]$ exists, $i \in \{b, p\}$.

b) Scenario #2: equilibrium in the presence of capacity payments

In the presence of capacity compensation payments the regulator fixes $K^T$, such that $K^T = k^*_{f,m} + k^*_{d,m} + k^*_{f,m} + k^*_{d,m}$, where $k^*_{f,m}$ is type-$i$ fringe generators’ equilibrium aggregate capacity and $k^*_{d,m}$ is type-$i$ dominant generator’s equilibrium capacity given a compensation for investment, $m$. The electricity generators that invest in capacity receive such a compensation $m$ per unit of capacity built. Having said that, the following proposition captures the competitive fringes’ aggregate equilibrium capacities.

**Proposition 13.** Equilibrium capacities chosen by the competitive fringe in the presence of capacity payments. Denote $a_i \equiv \bar{p} - c_i$, $i \in \{b, p\}$. Given $K^T \in (0, 1]$, at a interior solution, the base load fringe’s equilibrium aggregate capacity is $k^*_{f,m} = F^{-1} \left( F(K^T) - \frac{c_b - c_p}{c_p - c_b} \right)$ and the peak load fringe’s equilibrium aggregate capacity is $k^*_{f,m} = F^{-1} \left( F(K^T) - \frac{c_p - m}{a_p} \right) - F^{-1} \left( F(K^T) - \frac{c_b - c_k}{c_p - c_b} \right) - k^*_d$.

Likewise, fringes’ equilibrium aggregate investment in capacity is anticipated by the dominant firms. Therefore, a dominant generator $(i, d)$ chooses the period-2 profit-maximizing level of investment in capacity anticipating the fringes’ capacities and given both $(i, d)$ dominant generator’s investment and the capacity target level imposed by the regulator $K^T$. That is, for generator $(i, d)$, its optimal capacity investment solves:

$$k^*_{i,d,m} = \arg \max_{k^*_{i,d,m} \in [0, 1]} E\pi^d_i(p^*_s; p^I_i(\cdot), \theta)$$  \hfill (3.6)
where $E$ is the expected operator at period 1 and $\pi_{i}^{d,m}(p^{s}, p_{i}^{d}(\cdot), \theta)$ is generator $(i,d)$’s profit in the presence of a capacity compensation mechanism. The optimal choice of investment in generation capacities determines the equilibrium per-unit capacity compensation payment, denoted $m$.

**Proposition 14.** Dominant generators’ equilibrium capacities in the absence of capacity payments. $k_{i}^{*,d,m} \in [0,1]$ exists, $i \in \{b,p\}$. Moreover, the equilibrium compensation, $m$, is such $m = c_{k_{p}} + a_{p}[F(K^{T} - k_{p}^{*,d,m}) - F(K^{T})]$.

As in the competitive setting, in the dominant-fringe framework peak generators’ per-unit capacity costs determine the equilibrium capacity compensation mechanism.

### 3.6 Welfare analysis

As we mention in the introduction, the main goal of this chapter is to evaluate the desirability of a capacity adequacy requirement and capacity payments for consumers.\(^{39}\) For that purpose, we are interested in evaluating consumers’ welfare using the previously described equilibria. In particular, we compare the benchmark equilibrium case in terms of welfare with the equilibria obtained in the presence of a capacity compensation system. As discussed above, we focus our analysis on the interior solution cases, i.e. on the cases in which there is more than one production technology available in the market.\(^{40}\)

As is standard in the economics literature in general—and in the electricity markets literature in particular—we measure consumers’ welfare in terms of the (ex-ante)\(^{39}\) As Abito (2012) do in a similar context—along with the papers cited not in the next but in the following footnote—we assume that the social planner only cares about consumer welfare and thus, we put aside producer welfare. In the context of a competitive market, as we explain in the previous chapter, producer surplus loses relevance. In the context of a “dominant-fringe” market, producer surplus reflects the existence of market power which, as expected, leads to an inefficient outcome (typically in detriment of consumers). In addition, this assumption is not critical for the results. As this author points out, it suffices to have a social planner that puts a higher weight on the consumer welfare relative to generators welfare.

\(^{40}\)This accords with the reality that when a wholesale auction is cleared all the electricity is not produced using only one technology.
consumer surplus, i.e. as the price the consumers are willing to pay to get electricity (VOLL) minus the (expected) price paid by consumers multiplied by the total amount sold.\textsuperscript{41} In the absence of capacity payments, the expression for consumer surplus is given by:

\[
CS = \int_0^{k_b} (p^H - c_b)\theta dF(\theta) + \int_{k_b}^{K} (p^H - c_p)\theta dF(\theta) + \int_{K}^{1} (\max\{p^H - p^{\text{cap}}, 0\}) K dF(\theta)
\]

(3.7)

In the presence of capacity payments, consumer surplus is given by:

\[
CS^m = \int_0^{k_b} (p^H - c_b)\theta dF(\theta) + \int_{k_b}^{K} (p^H - c_p)\theta dF(\theta) + \int_{K}^{1} (\max\{p^H - p^{\text{cap}}, 0\}) K dF(\theta) - m
\]

(3.8)

where \(m\) is the per-unit capacity compensation payment that is passed-through to consumers as a charge paid in both peak and off-peak periods.

Notice that the consumer surplus expressions depend on two key components. First, they depend on the behavior of electricity prices. In particular, since consumer surplus depends on the difference between the maximum price that consumers are willing to pay and the price they actually pay, an increase in the expected price decreases welfare. Second, consumer surplus also depends on equilibrium capacity choices. In particular, an increase in capacity reduces the probability that the total installed generation capacity is not enough to serve the whole demand, i.e. it reduces the probability of quantitative electricity rationing (“blackouts”) and thus makes consumers (\textit{ex-ante}) better-off.\textsuperscript{42} Moreover, an increase in capacity reduces the probability of scarcity events and, therefore, reduces the probability of reaching either the VOLL \((p^H)\) or the price cap \((p^{\text{cap}})\). In other words, it reduces the probability of price-spikes and price volatility.

\textsuperscript{41}For the particular case of power markets, Allcott (2011), Borenstein (2005), Hartley and Kyle (1989), Newbery (1995), Sauma and Oren (2009) and White et al. (1996) (among many others) use the consumer surplus as the primary measure of welfare.

\textsuperscript{42}The fact that households have negative welfare effects from outages is the core of the discussion in Carlsson and Martinsson (2007), Carlsson and Martinsson (2008) and De Nooij et al. (2007). A similar analytical specification to ours for the “blackout” probability is in Brennan (2004).
Therefore, to analyze the welfare implications for consumers, it is also useful to consider the following expressions. First, we consider the expected amount paid by consumers, given by

\[ E(\text{payment}) = \int_{k_b}^{k_b} c_b \theta dF(\theta) + \int_{k_b}^{K} c_p \theta dF(\theta) + \int_{K}^{1} \min\{p^H, p^{\text{cap}}\} K dF(\theta) \quad (3.9) \]

if there is no capacity market and

\[ E(\text{payment}^m) = \int_{0}^{k_b} c_b \theta dF(\theta) + \int_{k_b}^{K_T} c_p \theta dF(\theta) + \int_{K_T}^{1} \min\{p^H, p^{\text{cap}}\} K_T dF(\theta) + m \quad (3.10) \]

if there is a capacity market. Second, we also consider the volatility of the price, calculated as the variance of the price, \( \text{vol(\text{price})} = \sum_s E(\text{price}^2) - E(\text{price})^2 \). Finally, we also report the probability of a “blackout”, defined as the probability that generation capacity is less than the demand, i.e. \( \text{pr}(K < \theta) \).

Since there are no general results implying that one equilibrium is always superior than another, in the next section we use data from the Texas ERCOT market to study different counterfactual scenarios.

3.7 A case study of the Texas ERCOT market

3.7.1 Background

The Electric Reliability Council of Texas (ERCOT), created in the seventies and established in 1996 as the first independent system operator (ISO) in the US, manages 85% of Texas’ electric load. In 2011 it served around 23 million customers, and had a total installed capacity of 84,000 MW. Generation was based on two main fuels: natural gas and coal. In fact, in 2011, 80% of the installed capacity in ERCOT was

\footnote{The assumed “rotated L” demand does not correspond to a simple utility function. Therefore, in our case, we are not able to calculate the change in expected utility.}
either coal (23%) or natural gas (57%). Wind, Nuclear, Biomass, Hydropower and other technologies are still residual in ERCOT.\footnote{In the numerical analysis presented in the following sections, we employ load data net of wind and nuclear. By doing so, we end up with the residual demand faced by coal and gas plants.}

Although ERCOT aims to ensure reliable power operations in a competitive wholesale market\footnote{According to the official website “ERCOT works with the electric industry organizations in the ERCOT control area to ensure reliable power operations for the wholesale and retail competitive markets”. See \url{http://www.ercot.com/mktparticipants/} [last access: 03/10/2016].}, the competitiveness of this market has been questioned –see, for instance, Dyer (2011). The Texas ERCOT market consists of a wide variety of investors: publicly-traded companies, private firms, municipally-owned utilities, cooperatives, etc., many of which operate a small amount of capacity. However, as pointed out by Hortacsu and Puller (2008), two major players, TXU and Reliant, controlled 24% and 18% of the total installed capacity respectively at the time they were writing. More recent data –see Newell et al. (2012)– shows that in 2012 Luminant (formerly named TXU) still controlled 17% of the total ERCOT generation capacity while NRG (formerly named Reliant) controlled 14%. This fact motivated the alternative approach of a market with two dominant firms facing a competitive fringe that is analyzed in Section 3.5.\footnote{Both Luminant (TXU) and NRG (Reliant) own both coal and natural gas generation facilities. In particular, according to their official websites, 52% of Luminant’s generation capacity is coal fired and 33% uses natural gas (see \url{http://www.luminant.com/plants/generation.aspx} [last access: 03/10/2016]). Data on NRG generation assets can be found at \url{http://www.nrg.com/about/who-we-are/our-assets/} [last access: 03/10/2016].}

As we mention in the introduction, Texas ERCOT relies on an “energy-only” market system. In other words, electricity prices adjust to equate demand and supply. If there is a shortage, prices rise above the marginal production cost and generate a scarcity rent. This scarcity rent is needed to recover investment costs. Capacity compensation payments are absent in the system.

Although the Texas ERCOT market appeared to be a successful case of a deregulated “energy-only” market, in 2010 and 2011 several reports warned that the system was not providing the right incentives to build new generation capacity. In fact, as
reflected in Figure 3.1, the capacity reserve margin, far from reaching the 13.75% target level, was shrinking and was projected to become negative.\footnote{The capacity reserve margin is calculated as follows: \[
\text{Capacity Reserve Margin} = \frac{\text{Generation Capacity Resources (MW)} - \text{Firm Load Forecast (MW)}}{\text{Firm Load Forecast (MW)}} \times 100
\]}

Figure 3.1: ERCOT forecasted reserve margin (2011)

In an attempt to solve this problem, in 2012 policymakers raised the System-Wide Offer Cap (SW-CAP), which is the highest per-MWh price at which a generation resource can offer energy in the market. In our model, the SW-CAP is the price cap \( (\bar{p}^\text{cap}) \). By doing so they expected to generate greater scarcity rents and, consequently, to incentivize new investment in generation capacity. Thus, as shown in Table 3.1, the price cap in ERCOT experienced a steady increase from $2,500/MWh in 2011 to $9,000/MWh in 2015. In fact, ERCOT’s long-term price cap goal was $15,000/MWh.\footnote{However, as discussed in Bajo-Buenestado (2015), this price cap raise created a problem of price-spikes in the Texas ERCOT market.} Thus, from 2011 on, ERCOT’s price cap was set well above the price cap in other markets such that the NYISO ($1,000/MWh), PJM ($1,000/MWh), or ISONE ($1,000/MWh). Nevertheless, questions continued to be asked about the reliability of the ER-
Table 3.1: ERCOT System-Wide Offer Cap (SW-CAP)

<table>
<thead>
<tr>
<th>Year</th>
<th>SW-CAP ($/MWh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011</td>
<td>2,500</td>
</tr>
<tr>
<td>2012</td>
<td>4,500</td>
</tr>
<tr>
<td>2013</td>
<td>5,000</td>
</tr>
<tr>
<td>2014</td>
<td>7,000</td>
</tr>
<tr>
<td>2015</td>
<td>9,000</td>
</tr>
<tr>
<td>Long-term goal</td>
<td>15,000</td>
</tr>
</tbody>
</table>

COT system beyond 2012. Thus, in June 2014 ERCOT implemented the so-called Operating Reserve Demand Curve (ORDC) system. The basic idea underlying this mechanism is that generators holding additional reserves are compensated whenever total system reserves cross a lower threshold. The compensation is immediately added to the wholesale price. As of the time of writing it is still unknown whether this change will solve the problem.

It seems that ERCOT members are reluctant to embrace the idea of a capacity market. In fact, although some authors claim that capacity markets have successfully solved similar capacity shortages in a range of different countries and regions –see Ausubel and Cramton (2010)– some other authors, such as Kleit and Michaels (2013a), Kleit and Michaels (2013b) and Kleit and Michaels (2013c), argue against the implementation of a capacity market in Texas. Therefore, and considering that most of the papers covering this debate are atheoretical, the question of whether or not a capacity market would be welfare enhancing for ERCOT consumers is still open.

3.7.2 Data

In this subsection, we use parameters and data from the Texas ERCOT market to simulate the equilibrium capacities –building on the results from Propositions 7, 8 and 9– and study welfare consequences of different counterfactual scenarios.

First, since installed capacity is dominated by both coal and natural gas plants, we assume that the former (coal plants) represent base load plants and the latter (natural
gas plants) represent peak load plants. Next, we assume the following parameters: \( p^H = 6000 \), \( p^{cap} = 2500 \), \( c_p = 42.03 \), \( c_b = 24.5 \), \( c_{kp} = 45.8 \) and \( c_{kb} = 63 \). The VOLL \( (p^H) \) is obtained from Frayer et al. (2013). According to this study, and based on 2011 macroeconomic indicators, the ERCOT-wide estimate of the VOLL is approximately \$6,000/MWh. The price cap \( (p^{cap}) \) value is the SW-CAP in the Texas ERCOT market in 2011, i.e. \$2,500/MWh. To obtain peak load generators’ variable cost \( (c_p) \) we multiply the heat rate for a natural gas plant, obtained from the EIA (8.185 MBtu/MWh) by the average cost of natural gas delivered for electricity generation in Texas, also obtained from the EIA (\$4.27/MBtu), and we add an estimate of the operations and maintenance (O&M) costs for natural gas plants in ERCOT, obtained from the ERCOT website (\$7.08/MWh). Thus, we reckon that the variable cost for peak load plants is \$42.03/MWh. To obtain the variable cost for base load generators \( (c_b) \) we similarly multiply the heat rate for coal plants (10.415 MBtu/MWh) by the average cost of coal delivered for electricity generation in Texas (\$1.87/MBtu), and we add the O&M costs (\$5.02/MWh). This gives a variable cost for base load plants of \$24.5/MWh.\(^{49}\) The per-unit capacity costs are assumed to be the levelized cost of capital for both a conventional gas-fired plant (peak) and a conventional coal plant (base). The levelized cost of capital represents the cost associated with the reimbursement of the initial capital investment of a generation plant. That is, it proxies the per-MWh capital cost of building a generating plant. The actual numbers that we use were obtained from the EIA’s Annual Energy Outlook 2011 (AEO2011).\(^{50}\)

Finally, we need a proxy for \( F(\theta) \), i.e. the cdf of the demand over the normalized support \([0, 1]\). For that purpose, we use ERCOT hourly load data (in MW) from 1996

\(^{49}\)All this information is publicly available. More precisely, the heat rates and the average cost of fuel delivered for electricity generation were obtained from the EIA website. The information about O&M costs was obtained from ERCOT website, and can be found in the document “396NPRR-01 Reduce Standard O&M Costs 071911”.

\(^{50}\)See http://www.eia.gov/oiaf/aeo/electricity_generation.html [last access: 03/22/2016].
to 2014.\footnote{Data is not available for the years 2001-2002.} Considering that investing in capacity is a long-term decision and that, in fact, the generating plants’ lifespans are extended up to a few decades, it makes sense to use an assessment of demand based on several years’ data to analyze investment decisions.

Furthermore, we add the following changes. First, we treat nuclear generation capacity as running 24/7 and subtract it from ERCOT’s total load. While not literally true, nuclear plants are only fuelled about once every 18 months and otherwise are very reliable. In fact, by treating them as running 24/7, and considering that negative load is considered as lost and thus normalized to zero, we get annual amounts of nuclear generation very close to the ones presented in ERCOT’s website.\footnote{There are two nuclear power plants in Texas, Comanche Peak Nuclear Power Plant and South Texas Nuclear Generating Station. The former, built in the early nineties, has a net generation capacity of 2,406MW after two expansions that took place in 2008 and 2009 and it is owned and operated by Luminant. The latter, built in the late eighties, has a total generation capacity of 2,560MW and it is 44% owned by NRG.} Second, we also subtract the wind power generation from the load and treat $\theta$ in the model as fluctuations in demand net of wind.\footnote{Hourly aggregated wind output is available since 2007. Before 2007, wind generation was marginal and almost non-existent.} Ignoring hydro and biomass, which represents the smallest proportion in ERCOT’s generation mix (less than 1% of the system’s generation capacity), we end up with residual demand faced by coal and gas plants.\footnote{For more information about the generation capacity mix see “ERCOT 2011 Quick Facts”, available at \url{http://www.ercot.com/content/news/presentations/2012/ERCOT%20Quick%20Facts%20-%20Jan2012.pdf} [last access: 03/22/2016].}

Using the aforementioned data (145,762 observations), we set the highest hourly demand (102,216.5 MW in August 6, 2003) equal to one, and then we normalize the rest of the hourly data relative to it. Next, with the normalized hourly load data, we estimate the kernel density of it, to obtain the pdf of the demand, $\hat{f}(\theta)$. That is, the estimated pdf at $\theta$ is given by

$$\hat{f}(\theta) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{\theta_i - \theta}{h}\right)$$ (3.11)
where $n = 145,762$ (observations), $K$ is the kernel function and $h$ is the bandwidth.\textsuperscript{55} The cdf, $\hat{F}(\theta)$, is just obtained by integrating $\hat{f}(\theta)$ over the desired interval. Figure 3.2 captures ERCOT’s load duration curve (net of wind and nuclear) from 1995 to 2014. The load duration curve, which is a common chart used in power markets, simply represents the inverse “survival function” of ERCOT’s load.

Figure 3.2: ERCOT hourly load duration curve, net of wind and nuclear generation (1996-2014)

\begin{center}
\includegraphics[width=0.5\textwidth]{load_duration_curve.png}
\end{center}

\subsection*{3.7.3 Numerical analysis and results}

Using the previous cost estimates, the estimated cdf of ERCOT hourly demand and the theoretical results in sections 3.3, 3.4 and 3.5, we now investigate the welfare consequences for consumers in Texas of different counterfactual scenarios.

\textbf{Perfect competition case}

First, putting aside the discussion of whether a capacity market is desirable for consumers or not, and given the aforementioned Texas ERCOT 2011 scarcity events, we

\textsuperscript{55}In this case, we use the (default) Gaussian kernel and the chosen bandwidth is the usual Silverman’s rule-of-thumb. See Silverman (1986).
analyze the impact on consumers of the steady increase in the price cap that ERCOT have been implementing since 2011. As we mentioned earlier, ERCOT’s price cap (SW-CAP) increased from $2,500/MWh in 2011 to $9,000/MWh in 2015. The main results of such a policy are presented in Table 3.2.

Table 3.2: Welfare gains of raising/eliminating the price cap (perfect competition)

<table>
<thead>
<tr>
<th></th>
<th>From $p_{\text{cap}} = 2500$ to $p_{\text{cap}} = 5000$</th>
<th>Eliminating the price cap $p_{\text{cap}} = 2500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Installed Capacity</td>
<td>+5.33%</td>
<td>+6.60%</td>
</tr>
<tr>
<td>Base load capacity</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Peak load capacity</td>
<td>+7.50%</td>
<td>+9.28%</td>
</tr>
<tr>
<td>Blackout probability</td>
<td>-0.009</td>
<td>-0.011</td>
</tr>
<tr>
<td>Price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected payment</td>
<td>+3.63%</td>
<td>+4.48%</td>
</tr>
<tr>
<td>Volatility</td>
<td>+42.74%</td>
<td>+56.59%</td>
</tr>
<tr>
<td>Consumer Surplus</td>
<td>+0.06%</td>
<td>+0.06%</td>
</tr>
</tbody>
</table>

All the values indicate either percentage change or change in probability.

On the one hand, a gradual increase in the price cap leads to a similarly gradual increase in the equilibrium level of investment in installed capacity. This is true until the price cap is not binding anymore. Once that happens, an extra increase in the price cap has no effect on generators’ investment in capacity since the VOLL will then be binding. This result is consistent not only with Corollary 4, but also with the actual evolution of ERCOT’s installed capacity. Even though the forecasted reserve margins in 2011 for the subsequent years were below 9%, in 2013 and 2014 –i.e. when the price cap was removed– the projected reserve margins were clearly above 13.75%, which is ERCOT reserve target. As the installed capacity increases, the reliability of the system improves. This is reflected in the middle row of Table 3.2, in which we can observe a steady decrease in the probability of a blackout due to insufficient generation capacity.

On the other hand, an increase in the price cap also increases the price volatility.
by increasing the likelihood of price spikes. Again, once the price cap is not binding, this effect disappears. Thus, raising the price cap may lead to events like the one experienced in September 3, 2013, when the market clearing price was $4,900/MWh—just $100/MWh below the price cap at the time—setting the new record in the State.\textsuperscript{56} In addition, an increase in the price cap also increases the expected amount paid by consumers. Such an increase goes up to 4.48% if the price cap disappears.

The overall effect on consumer surplus is captured in the bottom row. Thus, we can see that an increase in the price cap leads to a gradual increase in the consumers’ welfare, yielding a more efficient market outcome. Therefore, although price caps may have some beneficial effects in the sense of reducing politically unpopular price spikes—see Brown (2014)—and promoting cost reduction—see Cabral and Riordan (1991)—our result casts doubt on the benefits for consumers of introducing price caps, at least for the case of ERCOT in recent years.

Next, we investigate the effect of introducing a capacity market into the electricity sector. Specifically, we compare the equilibrium parameters in a market with and without capacity payments. We consider a scenario in which the capacity target is set to the highest possible level to avoid blackouts, i.e. $K^T = 1$. This scenario seems to be closer to actual capacity markets, since typically the target capacity is set equal to the expected peak load demand plus a reserve margin and to guarantee a supply shortage no more than once every ten years\textsuperscript{57}.

The main findings are displayed in Table 3.3, in which we provide the changes in the main consumers’ welfare parameters after introducing the capacity compensation mechanism.

Table 3.3 presents the results in the perfect competition case. As we see, the two...
Table 3.3: Welfare gains of introducing capacity payments (perfect competition)

<table>
<thead>
<tr>
<th></th>
<th>From cap ((p^\text{cap} = 2500)) to both cap ((p^\text{cap} = 2500)) &amp; capacity compensation</th>
<th>From no cap to both cap ((p^\text{cap} = 2500)) &amp; capacity compensation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Installed Capacity</td>
<td>+101.95%</td>
<td>+89.46%</td>
</tr>
<tr>
<td>Base load capacity</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Peak load capacity</td>
<td>+142.74%</td>
<td>+122.22%</td>
</tr>
<tr>
<td>Blackout probability</td>
<td>-0.02</td>
<td>-0.008</td>
</tr>
<tr>
<td>Price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected payment</td>
<td>+67.77%</td>
<td>+60.58%</td>
</tr>
<tr>
<td>Volatility</td>
<td>-99.33%</td>
<td>-99.57%</td>
</tr>
<tr>
<td>Consumer Surplus</td>
<td>-1.24%</td>
<td>-1.30%</td>
</tr>
</tbody>
</table>

All the values indicate either percentage change or change in probability.

cases presented yield similar results. The capacity compensation mechanism both increases system reliability and reduces price volatility. When there is a binding price cap both before and after, the probability of a blackout is reduced by 0.02. While a blackout is also less likely to occur when the price cap is not binding in the absence of capacity payments, the reduction in the probability of suffering it is only by 0.008. This result is also expected, as the regulator’s capacity target is set highly enough. Hence, a greater equilibrium investment in capacity reduces the probability of shortages and, consequently, the probability of price-spikes. This is again consistent with the data displayed in Table 3.3, where we observe a reduction in price volatility that ranges from 99.3% \((p^\text{cap} \text{ case})\) to 99.6% (not binding cap case).

However, these benefits come at a cost. In particular, the introduction of a capacity market substantially increases the expected amount paid by consumers. This is due to the fact that capacity payments are effectively passed-through to consumers. Overall, we find a negative impact on consumers of about -1.25%. In other words, the introduction of a capacity payment yields an inefficient outcome in the market.
**Imperfect competition case**

Again, we first analyze the impact on consumers of the increase in the price cap that ERCOT implemented since 2011 for a market with two dominant firms and a competitive fringe. The results are included in Table 3.4. In fact, the results are very similar to those obtained in the competitive market scenario. The only difference is that the marginal contribution of an increase in the price cap to the system reliability is lower. This is true because, in a market with a competitive fringe (for both base and peak load) and two dominant firms (i.e. with many firms), the installed capacity is expected to be high enough to avoid blackouts even during peak-demand periods. However, in contrast to the perfectly competitive case, consumer surplus decreases by -0.07%.

Table 3.4: Welfare gains of raising/eliminating the price cap (imperfect competition)

<table>
<thead>
<tr>
<th>From $p^{\text{cap}} = 2500$ to $p^{\text{cap}} = 5000$</th>
<th>Eliminating the price cap $p^{\text{cap}} = 2500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Installed Capacity</td>
<td>+2.53%</td>
</tr>
<tr>
<td>Base load capacity</td>
<td>+0.05%</td>
</tr>
<tr>
<td>Peak load capacity</td>
<td>+4.88%</td>
</tr>
<tr>
<td>Blackout probability</td>
<td>-0.0005</td>
</tr>
<tr>
<td>Price</td>
<td></td>
</tr>
<tr>
<td>Expected payment</td>
<td>+2.70%</td>
</tr>
<tr>
<td>Volatility</td>
<td>+42.74%</td>
</tr>
<tr>
<td>Consumer Surplus</td>
<td>-0.06%</td>
</tr>
</tbody>
</table>

All the values indicate either percentage change or change in probability

Finally, we investigate the effect of introducing a capacity market into the electricity sector. The main results are displayed in Table 3.5. Again, the introduction of capacity payments reduces both the probability of blackout (i.e. increases aggregate installed capacity) and price volatility for consumers. However, as indicated in fifth row, this policy measure increases substantially the expected amount paid by consumers. Overall, the impact on consumer surplus is negative in the two scenarios
included in the table.

Table 3.5: Welfare gains of introducing capacity payments (imperfect competition)

<table>
<thead>
<tr>
<th></th>
<th>From cap ((p_{\text{cap}} = 2500)) to both cap ((p_{\text{cap}} = 2500)) &amp; capacity compensation</th>
<th>From no cap to both cap ((p_{\text{cap}} = 2500)) &amp; capacity compensation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Installed Capacity</td>
<td>+72.89%</td>
<td>+67.99%</td>
</tr>
<tr>
<td></td>
<td>+128.97%</td>
<td>+128.86%</td>
</tr>
<tr>
<td></td>
<td>+19.94%</td>
<td>+13.55%</td>
</tr>
<tr>
<td>Blackout probability</td>
<td>-0.001</td>
<td>-0.0002</td>
</tr>
<tr>
<td>Price</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>+60.81%</td>
<td>+55.69%</td>
</tr>
<tr>
<td>Volatility</td>
<td>-91.87%</td>
<td>-94.81%</td>
</tr>
<tr>
<td>Consumer Surplus</td>
<td>-1.37%</td>
<td>-1.30%</td>
</tr>
</tbody>
</table>

All the values indicate either percentage change or change in probability

In summary, in the context of the Texas ERCOT market, we find that price caps are consumer-surplus enhancing if dominant firms are present in the market. Otherwise, i.e. if the market behaves competitively, price caps reduce efficiency. Regarding a capacity market, we find the following. On the one hand, a capacity market improves system reliability and reduces electricity prices oscillations. It does so, however, at the cost of increasing consumers’ expected payments. Overall, we find that a capacity market has a negative impact on consumer surplus.

### 3.8 Conclusion

In this chapter we investigate the welfare gains for consumers of introducing a capacity market in the context of a price capped power market. Our main findings are as follows. First, in line with Joskow and Tirole (2007), we cast doubt on the potential benefits of introducing price caps. In fact, our model indicates that their introduction leads to less investment in capacity, which implies a higher blackout probability to the detriment of consumers. However, in the presence of dominant generators in the
market, price caps are consumer-surplus enhancing. The reason is that they prevent the dominant players from raising prices and exercising market power.

Second, we conclude that a capacity market serves as an “insurance mechanism” for consumers. It reduces the likelihood of extreme events –i.e. prices spikes and blackouts– by increasing system reliability and reducing price volatility, at the cost of increasing consumers’ average expected payments. For different price caps, we find that consumer surplus goes down in all the cases in the example based on ERCOT data.

Our results were obtained with little or no assumptions about consumers’ attitudes toward risk. In other words, the benefits of introducing a capacity compensation mechanism, which are essentially due to reduction of extreme events like price spikes and blackouts, are achieved without assuming that consumers are risk averse. If we were to assume that, we expect that the welfare benefits of higher generating capacity would be even greater. We have also avoided the question of whether the capacity compensation mechanism should be implemented using an auction or via bilateral trading. Although this has also been a matter of intense debate, we believe that it is beyond the scope of this chapter. Further extensions should consider that issue. Another useful extension would involve taking the dynamic setting of the investment decision more seriously.
Part II

Retail Gasoline Markets
Chapter 4

Evidence of asymmetric behavioral responses to changes in gasoline prices and taxes for different fuel types

4.1 Introduction

The second half of the 2000s was a tough period for the public finances in several countries in the advanced world. The poor evolution of unemployment, GDP, savings, and/or private consumption –among other economic indicators– seriously compromised both the government revenue as well as the debt-to-GDP ratio of many of these countries.

Because of budget shortfalls, many governments chose to raise taxes. Gasoline was a common target. Indeed, over the last years, and coinciding with the onset of the economic downturn, many of the governments in trouble decided to increase taxes on gasoline.
That was true in the following countries (among many others):

- Lithuania and Latvia in 2009.

- Romania, first in 2009—see Box 1: “Overview of the main tax related measures taken in response to the economic and financial crisis” in European Commission (2009)—and again in 2014, when there was a 7 cents per litre increase in excise duties for unleaded petrol, leaded petrol, diesel and kerosene (used as motor fuels)—see European Commission (2014b).

- Greece, where the price of unleaded gasoline increased by 0.12 euro per litre and the price of diesel fuel increased by 0.05 euros a litre in February 2010 in order to mitigate the negative economic and budgetary impact of the global crisis—see European Commission (2010).

- Italy, where the tax incremented first by 0.005 euro per litre and then by 0.02 euro per litre with the goal of increasing the revenue to mitigate the 2009 and 2012 earthquakes damages respectively.

- Netherlands, where the excise duty on Liquefied Petroleum Gas (LPG) products went up by 7 cents per litre and the excise duty of diesel by 4 cents per litre in January 2014—see European Commission (2014a).


As the politicians in some of these countries indicated, there were also environmental and climate change objectives behind such policies. However, in most of the cases (if not all) the main goal behind these tax increases was to raise public revenues in a complicated situation for the public treasury.¹

This was the case, for example, in Spain. This Southern European country suffered after 2007 a progressive deterioration of its key macroeconomic indicators –especially the unemployment, which rose from approximately 10% in 2008 to around 25% in 2014, according to the OECD (2015) (p.106). At the same time, the government was running out of money. In fact, while the public debt was 35.5% of GDP in 2007, it increased to 92.1% of GDP by the end of 2013, and it is expected to reach a peak of 103.2% of GDP in 2017 according to the European Commission (2015) (p.34). Thus, following the advice of the European institutions and other international bodies –see ECB (2008), ECB (2009) and IMF (2007), among many others–, and in order to demonstrate fiscal discipline, Spain had to consolidate its public accounts by cutting the public spending and increasing the taxes. Thus, as many other countries’ governments did, the Spanish government decided to raise gasoline taxes for that purpose. Indeed the Spanish Council of Ministers approved on June 13, 2009 an increase of 0.029 euro per litre of the excise duties on diesel fuels and unleaded gasoline fuels. In fact the government recognized that the main goal behind such a measure was to increase tax collection in a very complicated situation for the Spanish public treasury\(^2\), but was this decision a wise one in terms of tax collection? Was the Spanish government –and the rest of the countries with similar policies– doing the right thing?

We want to point out that actually a raise in gasoline taxes is not as effective as previous studies have concluded. In particular, we cast doubt on previous studies that estimated the effect of taxes based on the “overall” (i.e. the tax-inclusive) elasticity of the gasoline demand. For that purpose, using detailed level data from Spain on different fuel types, we estimate consumers’ responses to changes in gasoline consumption by taking into account separately both changes in tax-exclusive gasoline prices and changes in gasoline taxes. We provide empirical evidence to state that, for

\(^2\)Real Decreto-ley 8/2009, de 12 de junio, por el que se conceden créditos extraordinarios y suplementos de crédito, por importe total de 19.821,28 millones de euros, y se modifican determinados preceptos de la Ley 38/1992, de 28 de diciembre, de Impuestos Especiales. BOE, June 13, 2009 no. 143, pp. 49890-49902.
the case of unleaded gasoline (both regular and premium) and for agricultural diesel, the price elasticity of demand due to changes in taxes is much higher than the price elasticity of the demand due to changes in the tax-exclusive price. In other words, we demonstrate that an increase in gasoline taxes implies a greater reduction in gasoline consumption than an equal-sized increase in gasoline “pre-tax” prices. We show that the results are robust to alternative specifications that take into account potential dynamic adjustments in the consumption patterns and we also validate this result using an instrument for the price of the fuels.

On the other hand, and contrary to the evidence that Li et al. (2014) found for the unleaded gasoline case and that we also refute for the agricultural diesel fuel, such asymmetric response is not found in the (regular) diesel case. In fact, we show in all the different specifications of our baseline model that the consumption of gasoline decreases equally due to increases in taxes and prices. This result is consistent with the fact that frequent drivers are usually “diesel drivers”, so they tend to be better informed about changes in both taxes and prices –see Baranzini and Weber (2013) and Verboven (2002).

The study of the price elasticity of gasoline demand is a classic topic in the energy economics literature. Indeed, many authors have previously studied and empirically assessed this elasticity in different countries, in different periods of time and using all kind of approaches. However, the literature has not explored as deeply the differences that arise when we estimate separately the impact of changes in taxes and in “pre-tax” prices.

---

3 Notice that the factors that impact the price of both unleaded gasoline and diesel fuel are quite similar. In particular, both types of fuels have a reference index in both the Genoa market (MED) and Rotterdam market (NWE), which are the relevant markets for the Spanish wholesale gasoline sector. As shown in Rodrigues (2009) (Chart 2 and Chart 3) the MED and NWE prices for both fuels evolve similarly to the Brent crude oil price. In addition, there is also a wholesalers’ markup, which is also similar for both fuels and is approximately 2 per cent of the final price according to the information provided by the Spanish Association of Operators of Oil Products (AOP).

4 For instance, Akinboade et al. (2008), Baranzini and Weber (2013), Brons et al. (2008), Galindo (2005), Havranek et al. (2012), Lin and Zeng (2013) and Ramanathan (1999) make up just a small sample of the previous papers that have studied the elasticity of gasoline demand.
The most relevant study investigating this issue is by Li et al. (2014), whose paper is closely related to this chapter. Thus, as they do, we are concerned about the existence of asymmetric behavioral responses in consumption due to changes in prices and taxes, and we seek to reinforce their conclusions. However, our approach differs from theirs in a few ways. First of all, we want to check that this asymmetric effect holds for all the main fuels used for transportation purposes. For that reason we estimate the elasticities not only for unleaded gasoline, but also for premium unleaded gasoline, regular diesel fuel as well as agricultural diesel fuel, which is mainly used by tractors. By doing so, we are able to identify that asymmetric responses are not equally observed in all fuel types. In fact, for the Spanish case, it does not hold for regular diesel fuel consumers.

Second, instead of using annual data, we use monthly data. Thus, as Klier and Linn (2010) do, we are able to check that this effect is not only true in the long-run, but also in the short-run.\textsuperscript{5} Another closely related work is by Davis and Kilian (2011). They also explicitly recognize that “the responsiveness of gasoline consumption [due] to a change in tax may differ from the responsiveness of consumption [due] to an average change in price”. However, they take this fact into account to perform a different kind of analysis, namely, they explore the potential impact of a carbon tax in carbon emissions.

Taking the advantage that almost all the regional governments in Spain implemented several changes in excise duties on gasoline after the crisis—along with the few changes implemented by the central government—we find evidence to support the fact that, at least in Spain, the sensitivity of gasoline consumption to changes in taxes is greater than the sensitivity of gasoline consumption due to changes in tax-exclusive prices for three fuel types, namely unleaded gasoline 95, unleaded gasoline 98 (premium) and diesel B (agricultural). As a robustness check, we also perform

\textsuperscript{5}Notice that Li et al. (2014) use a monthly model as part of the robustness analysis (see Table 5). However, no control variables are included in such a model.
an analysis taking into account potential dynamics effects, namely, a lagged effect of (agricultural) unemployment on diesel B consumption\(^6\) as well as the existence of a partial adjustment in the consumption of gasoline. Moreover, in Section 4.2.4, and due to the potential concern of endogeneity of fuel prices and fuel consumption, we propose an instrumental variable regression. Again, in this regression, the asymmetric behavior is also refuted.

Previous literature has given (at least) up to two different (and complementary) potential explanations of such an asymmetric effect. First of all, following Li et al. (2014) these asymmetric responses may be due to the fact that there is a persistent effect in gasoline consumption. In other words, since cars are durable goods, consumers’ responses depend not only on today’s price of gasoline, but also on expected future gasoline prices. Thus, if consumers assess that a change in taxes is more persistent than a change in gasoline “pre-tax” prices, they should react by buying less motor vehicles and/or driving less in the former case than in the latter.

Second, some other authors—for instance, Chetty et al. (2009) and Rivers and Schaufele (2012) among others—have argued that there is a problem of “salience” with taxes. In other words, it is much more costly to calculate the impact of daily changes in “pre-tax” prices on the gross-of-tax price of gasoline than the impact of a single and publicly observable change in the taxation of gasoline on the gross-of-tax price of gasoline. Therefore, we should expect that a change in taxes has a greater negative impact on the demand—it is more salient—than a similar change in the tax-exclusive price.

This second explanation is in line with recent and innovative literature on industrial organization and behavioral economics. Thus, as Spiegler (2011) indicated in the context of gasoline consumption, consumers obtain information about gasoline prices using a “sampling” procedure. In other words, consumers elaborate expecta-

\(^{6}\)Jimeno and Bentolila (1998) find evidence that the impact of unemployment on wages is lagged. Therefore, we could expect that there is also a lagged effect on gasoline consumption too.
tions about gasoline prices using the prices they see in a subset of gas stations that they frequently pass by –gas stations that are close to their homes, those that are on the way to their works, etc.– from time to time. Thus, small daily changes in prices are negligible to them or not taken into account before the sampling procedure is repeated again. On the other hand, since an increase in the excise tax has media diffusion, they can automatically adjust this change in price at the moment the tax is introduced.\footnote{This effect, together with many others, is included in a comprehensive study on behavioral economics and taxes by Congdon et al. (2009).}

We can also find some explanations in previous literature to the fact that these asymmetric responses are not observed in diesel. As Schipper et al. (2002) note, diesel cars tend to be used more for commercial purposes (e.g. taxi drivers), and as private vehicles they tend to be bought by relatively high-kilometer drivers in Europe. Thus, we expect that diesel drivers (commercial users and frequent drivers) are usually better informed about changes in both prices and taxes, so they will react in a similar way to (similar) changes in taxes and prices.

There are several implications of these findings in terms of energy policy for the particular case of Spain. First of all, it seems that taxing at the pump is not as effective as previously thought in terms of revenue collection, especially if the taxes are levied on unleaded gasoline consumers. Thus, when we take into account separately changes in taxes and changes in tax-exclusive prices, we see that while the latter do not have a significant impact on gasoline demand, the former lead to a substantial decrease in gasoline consumption. Indeed this effect has been ignored by previous literature, leading to erroneous conclusions of the effect of a tax increase.

Therefore, we conjecture that whenever the goal of the government is to increase the tax revenue, it might be better to implement a different tax rather than a “tax at the pump” –see Goulder (1994) for a deeper discussion. For instance, it could be better to increase the taxes on gas stations’ profits or revenues: if the tax is effec-
tively pass-through to consumers\(^8\) via progressively higher prices, and since changes in prices are more unnoticeable for them, this measure can lead to a higher tax revenue. Therefore, a gas stations’ profits/revenues tax may be less distortionary than a “tax at the pump”.

This evidence also leads us to another conclusion in terms of climate change and environmental goals. In particular, since increases in taxes discourage gasoline consumption much more than increases in “pre-tax” prices, it is possible that previous analysis on environmental gains due to raises in taxes underestimate the benefits of these taxes in terms of reducing pollution and quantification of the climate change effect.

Second, notice that in many cases, policy makers implement the same increase in the tax rate for both diesel fuels and unleaded gasoline (for instance, in the aforementioned Spanish case, there was an increase of 0.029 euro per litre of the excise duties on diesel fuels and unleaded gasoline fuels). However, considering that the responsiveness of tax changes is different in diesel and unleaded gasoline consumption, it might be also a good idea for the government to implement different rates for different fuels. Thus, in the same way that car manufacturers price discriminate when selling diesel and unleaded cars, policy makers should “tax discriminate” (i.e. implement different tax rates for different fuels) when implementing new policies—see Verboven (2002).

The rest of the chapter is organized as follows. Section 4.2 provides the baseline empirical model and the robustness checks that we consider, describes the data, and provides summary statistics for the taxes by fuel type at the regional level. Section 4.3 includes the main results of the specified models and a brief discussion about them. Section 4.4 concludes and provides some policy implications.

\(^8\)Marion and Muehlegger (2011) find that gasoline taxes are indeed fully passed onto consumers and are incorporated fully into the tax-inclusive price, under typical supply and demand conditions.
4.2 Methods and data

In this section we provide the baseline model that we use to examine the elasticity of demand due to changes in tax-exclusive prices and taxes, as well as the alternative models used as robustness checks. We also give detailed information about the data we use to examine such models. We also provide evidence of the variability of the main variables that we use, namely fuel consumption, “pre-tax” prices and taxes.\footnote{In addition to that, in the Appendix (Table C.2) we provide summary statistics for the (nominal) taxes per fuel type during the examined period at the regional level due to the potential concern that taxes did not change very often in the examined period (indeed, as we can see, most of the regional governments introduced different changes).}

4.2.1 The baseline model

The hypothesis we want to test in this chapter is the following.

**Hypothesis 1.** An increase in fuel taxes implies a greater reduction in fuel consumption than an equal-sized increase in tax-exclusive fuel prices.

The model that we examine in this chapter is very similar to the ones used in previous studies on fuel demand –see, for instance, Hughes et al. (2008) and Li et al. (2014). In our basic set up –as well as in the proposed extensions–, the unit of observation is a province $i$ and month $t$. In the basic framework, our regression for each type of fuel, $G$, is given by:

$$
\log D_{it}^G = \beta_0 + \beta_1 \log p_{it}^G + \beta_2 \log \tau_{it}^G + \Theta X_{it} + \alpha_i + \alpha_t + \varepsilon_{it}^G \quad (4.1)
$$

where $D_{it}^G$ is fuel $G$ consumption in month $t$ in province $i$; $p_{it}^G$ is the average monthly retail price before tax of fuel $G$ in euro per litre ($\text{€/l}$) in month $t$ in province $i$; $\tau_{it}^G$ is the average tax of fuel $G$ in euro per litre ($\text{€/l}$) in month $t$ in province $i$. The scalars $\alpha_i$ and $\alpha_t$ denote province and month fixed effects, respectively. In addition, as Li et al. (2014) and Small and Van Dender (2007) do, we also include in the model...
a vector of control variables, $X_{it}$, namely the number of registered cars, trucks and vans, buses, motorcycles, tractors and other vehicles powered by fuel $G$ in month $t$ in province $i$.\textsuperscript{10} In section 4.2.2 some other controls that account for potential dynamic effects are also discussed.

One potential concern of this specification is that, unlike Dahl (2012), Hughes et al. (2008) and Li et al. (2014) do, we do not include income in the set of control variables. The main reason is that there is no monthly data on income available at the province level in Spain. However, the potential omitted variable bias is remedied by including two additional controls that are expected to be positively correlated with income.

First, we include as a control variable the amount of euros spent (in real terms) in house purchases per province per month. Considering that Spain is a country in which house rental is relatively low\textsuperscript{11}, there is an extensive literature proving the positive relationship between housing expenditure and income not only in Spain but in many other countries –see, among many others, Fernández-Kranz and Hon (2006), Gallin (2006), Garcia and Raya (2011), Lopez et al. (1998) and Martinez and Maza (2003). Moreover, we believe this variable captures the regional differences in economic activities and cost of living –see Garrido-Yserte et al. (2012) and Martinez and Maza (2003)– which are determinants of the income level. Second, we consider monthly data on consumer credits (in real terms) per province and month to measure consumer expenditure which, as expected, proxies consumer disposable income.\textsuperscript{12}

Finally, for the case of the agricultural diesel fuel, we also include as a control variable the agricultural unemployment level. By introducing this variable we take

\textsuperscript{10}However, unlike the aforementioned authors, we do not include some other demographic variables, such as the share of population living in rural areas or the number of drivers. The are two reason for that. First, because we do not have demographic data at the month level. Second, because we expect to see none or little variation of these variables in the monthly data.

\textsuperscript{11}As it is indicated by Ortega et al. (2011), “One of the most salient feature of the Spanish housing market, compared to other European economies, is its relatively low rental share.”

\textsuperscript{12}For the Spanish case, Pardo and Sánchez Santos (2014) provide some evidence of a positive link between household debt level and income after 2005.
into account the impact of changes in agricultural production in the consumption of diesel B.\textsuperscript{13}

A second potential concern about this baseline model is the issue of endogeneity (simultaneity) between fuel prices and fuel consumption. As Hughes et al. (2008) pointed out (among many others), since prices are potentially correlated with demand shocks, we might expect the estimated price elasticities of the demand to be downward biased, and thus underestimated.\textsuperscript{14} On the other hand, taxes are presumably less correlated with the demand, since they are not endogenously determined.

To address this potential endogeneity problem, we use instruments for the prices. In particular, we use the Brent crude oil price as an instrumental variable in our analysis. The reason for that relies on the fact that demand shocks at the province level are not expected to be correlated with crude oil prices (Brent), since these prices are set in the global oil market. Therefore, we expect that Brent is correlated with the gasoline prices but not with unobserved province-level gasoline demand shocks. Moreover, since the price of crude oil is the same across provinces for a given month, we interact the variable Brent with a province specific variable, in a similar fashion as Acemoglu et al. (2013) do. In particular, we use the CPI index by province. The reason to do so is because we expect higher gasoline prices in those provinces in which overall prices are higher. However, we do not expect strong correlation between gasoline demand shocks and CPI.

\textsuperscript{13}Considering that Villaverde and Maza (2009) find robust evidence that Okun’s law holds in Spain both at the national level as well as at the regional level, it is possible to think that unemployment should be included as a control for all the other fuels. However, the introduction of this variable for the unleaded fuel cases and for diesel A case yields notorious problems, such as collinearity and potential measurement error bias.

\textsuperscript{14}In the context of the US gasoline retail market, Li et al. (2014) also indicate that “the tax-exclusive gasoline price may not be exogenous to gasoline consumption if [one thinks that] state-level demand or supply shocks are correlated with equilibrium prices and consumption”.

4.2.2 Robustness check I: partial adjustment models in gasoline demand

Following the idea suggested by Houthakker et al. (1974) – and as Baltagi et al. (2003), Hughes et al. (2008) and Pock (2007) do, among many others – we analyze a similar two-way fixed effects regression in which we include gasoline consumption in the previous month, instead of contemporaneous gasoline consumption. The main reason for that is that, as Hughes et al. (2008) point out, “frictions in the market may prevent reaching the appropriate equilibrium level and, as a result, only a fraction of the desired change in consumption between [two months] is realized”. Therefore, we denote the equilibrium demand of gasoline as \( D_{it}^G^* \) and following Houthakker et al. (1974) we define the adjustment equation as:

\[
\frac{D_{it}^G}{D_{it-1}^G} = \left( \frac{D_{it}^G^*}{D_{it-1}^G} \right)^\lambda
\]  

(4.2)

for some \( \lambda \in (0, 1) \).

Next, reconsider equation 4.1. If we substitute the equilibrium demand of gasoline as the dependent variable, we get the following expression:

\[
\log D_{it}^G^* = \beta_0 + \beta_1 \log p_{it}^G + \beta_2 \log \tau_{it}^G + \Theta X_{it}
\]  

(4.3)

Thus, substituting equation 4.2 into our equation 4.3:

\[
\log D_{it}^G = \lambda \beta_0 + (1 - \lambda) \log D_{it-1}^G + \lambda \beta_1 \log p_{it}^G + \lambda \beta_2 \log \tau_{it}^G + \lambda \Theta X_{it}
\]  

(4.4)

Renaming the coefficients, we arrive at:

\[
\log D_{it}^G = \delta_0 + \delta_1 \log D_{it-1}^G + \delta_2 \log p_{it}^G + \delta_3 \log \tau_{it}^G + \Theta' X_{it}
\]  

(4.5)
where \( \delta_0 = \lambda \beta_0, \delta_1 = 1 - \lambda, \delta_2 = \lambda \beta_1, \delta_3 = \lambda \beta_2 \) and \( \lambda \Theta = \Theta' \).

### 4.2.3 Robustness check II: lagged effect of unemployment on diesel B consumption

As a second robustness check, we include in our diesel B regression not contemporaneous changes in agricultural unemployment, but previous-period changes on agricultural unemployment. The reason for that is that, as Jimeno and Bentolila (1998) find in their seminal paper, there is evidence in Spain that the effect of unemployment on wages is not immediate but rather lagged. Therefore we expect that, likewise, the effect of agricultural unemployment on diesel B consumption is not contemporaneous but rather lagged.\(^\text{15}\) Thus, we propose the following alternative model:

\[
\log(D^G_{it}) = \gamma_0 + \gamma_1 \log u_{it-1} + \gamma_2 \log p^G_{it} + \gamma_3 \log \tau^G_{it} + \Theta'' X_{it} \tag{4.6}
\]

where \( u_{it-1} \) is the agricultural unemployment in the previous month \((t - 1)\).

### 4.2.4 Robustness check III: instrumental variable regression

As we discuss above, some authors have raised the issue of potential endogeneity between gasoline prices and gasoline demand. Since prices are determined in equilibrium, this fact may lead to a potential downward bias of our estimator of the fuel price –see Coglianese et al. (2015) for an in-depth discussion on the use of instruments in this context.

For that reason, we estimate the same regression specified in equation 4.1, but we instrument fuel prices with crude oil prices. In particular, we use the price of Brent crude oil in euro per litre as an instrument. Since Brent prices are set in equilibrium

\(^{15}\)One might think that the effect of unemployment is cumulative. Thus, as an additional robustness check, we include both the agricultural unemployment in the previous month and the agricultural unemployment in the current month.
in the global oil market, we expect that the demand shocks of gasoline at the province level are not correlated with the Brent variable. However, the price of Brent is the same for all provinces for a given month. Therefore, we interact it with province level data. In particular, we interact the price of crude oil with monthly data on CPI per province in Spain, which is also presumably highly correlated with gasoline prices, but not with gasoline demand.

4.2.5 Data and summary statistics

In our analysis we use Spanish monthly data per province and per type of fuel. Data on fuel prices, fuel consumption and taxes is available from January 2011 to October 2014. There are 50 provinces in Spain, plus Ceuta and Melilla, which are considered separated “provinces”. In all the model specifications we study four types of fuel: diesel A, diesel B, unleaded gasoline 95 RON\(^{16}\) and unleaded gasoline 98 RON.\(^{17}\) Diesel A and unleaded gasoline 95 are the two most commonly used fuels for cars, motorcycles and small trucks and vans; diesel B is used by agricultural vehicles (tractors); and unleaded gasoline 98 is the premium brand of unleaded gasoline 95. Consumption data is given in metric tons per month, province and fuel type. However, data for diesel B (agricultural) consumption is missing for the Canary Islands.

Prices are given in euro per litre (€/l) and are reported as average prices per month, province and fuel type. The data set contains both “pre-tax” prices and “post-tax” prices. Tax-inclusive prices are calculated as the arithmetic average of the taxes of each gas station per province, month and fuel type. Hence, the taxes are obtained as the difference between tax-inclusive prices and tax-exclusive prices. According to the information provided by the Comisión Nacional de los Mercados y de la Competencia (CNMC) –which is the source of these variables–, there are three

\(^{16}\)Research Octane Number.

\(^{17}\)Data for other fuels is also available –for instance, biodiesel or “special-purposes diesel” (NGO, or Nuevos Gasóleos). However, their consumption is still marginal.
taxes included in the “post-tax” prices:\textsuperscript{18} the Value-Added Tax (VAT)\textsuperscript{19}, the excise tax\textsuperscript{20} and the (retail sales) gasoline tax.\textsuperscript{21} The first one (VAT) is imposed by the central government, so it is actually the same for all the provinces.

Between 2011 and 2014 there was one modification of the VAT: in September 2012, the general\textsuperscript{22} rate increased from 18\% to 21\%. On the other hand, the provinces have different excise tax and gasoline sales tax rates.\textsuperscript{23} These rates are actually decided by the governments of the different autonomous communities of Spain.\textsuperscript{24,25} During the period of our study, several changes were introduced in most of the autonomous communities. To check this, we include in the Appendix Table C.2, which includes the main summary statistics for nominal taxes per type of fuel by autonomous community. There are just two regions with no variation in taxes at all: Ceuta and Melilla. The reason is that these regions, which are considered “overseas territories”, have different tax regimes\textsuperscript{26}, so neither the VAT nor the excise duties are applied there. Hence we remove the observations for Ceuta and Melilla from our dataset.

Furthermore, we use monthly Consumer Price Index (CPI) data per province to express both tax-exclusive prices as well as taxes in real terms, using January 2011 as the base period. To obtain the real values we use the standard formula provided,

\textsuperscript{18}The three of them are directly charged to the consumers.
\textsuperscript{19}Impuesto sobre el Valor Añadido (IVA).
\textsuperscript{20}Impuesto Especiales (IIEE).
\textsuperscript{21}Impuesto de Venta Minorista de Determinados Hidrocarburos (IVMDH).
\textsuperscript{22}The general rate is the one applicable to fuel consumption.
\textsuperscript{23}As we have indicated in the introduction, there is also an excise duty on gasoline that depends on the central government.
\textsuperscript{24}There are 17 autonomous communities plus Ceuta and Melilla, which are considered as two different “Autonomous Communities”. Nine of these autonomous communities have just one province. Two have two provinces. Three have three provinces. Two have four provinces. One has five provinces. One has eight provinces. And one has nine provinces.
\textsuperscript{25}Some small differences are also present in different provinces within an autonomous community. These differences are due to minor subsidies at the municipal level. However, these policies are infrequent and thus negligible.
\textsuperscript{26}For instance, there is not VAT.
for instance, by the EIA (Energy Information Administration).\footnote{To obtain the real prices and taxes, we apply the following standard formula: \[ \text{Real Price}_{i,t} = \text{Nominal Price}_{i,t} \frac{\text{CPI}}{\text{CPI}_{i,\text{January 2011}}} \]} Data on CPI is publicly available in the Spanish National Statistics Institute (INE) website.

Both data on consumption and prices per month, province and fuel type were obtained from the aforementioned CNMC, which is a public organization of the Spanish Ministry of Economy and Finance. The data is publicly available. Data on vehicles was obtained from the Dirección General de Tráfico (DGT), which is a public organization of the Spanish Ministry of the Interior. In particular, we use the total number of registered vehicles per province, month and type of fuel. The dataset divides the vehicles in the following categories: cars, vans (including Light or Medium Duty Trucks), buses, motorcycles, tractors and other vehicles.\footnote{The dataset also contains data on trailers and semi-trailers, but in many provinces the number of trailers and semi-trailers is zero, while in the rest of the provinces the number of trailer is very small and insignificant.} The data is available until September 2014 and is publicly available.

Data on consumer credits and on the amount spent on house purchases was obtained from the Consejo General del Notariado (CGN), which is a public organization of the Spanish Ministry of Justice. The data is available from January 2011 until October 2014 and it is publicly downloadable. Both variables are expressed in real terms (see footnote 27). Data on agricultural unemployment was obtained from the Servicio Público de Empleo Estatal (SEPE), which is a public organization of the Spanish Ministry of Employment and Social Security. In particular, we use the total number of registered unemployed in the agricultural sector per province and month. The data is available for the whole period of the study –i.e. from January 2011 to October 2014–, and is publicly available and downloadable.

Finally, for the proposed instrument we use several sources. Data for the price of Brent is from the EIA. In particular, we use monthly Europe Brent Spot Price FOB,
in dollars per barrel. We use data on the euro/dollar exchange rate provided by the Federal Reserve Bank of St. Louis to express the price of Brent in euros per litre.  

Table 4.1 summarizes the extent of data inter and intra-province variation for the main variables. Both the variation of fuel consumption as well as the variation of taxes (except for diesel B) are predominantly between province. However, the tax-exclusive prices variation is mostly within province. This fact might lead us to think that the use of heterogeneous estimators are convenient to capture the effect of tax-exclusive prices on fuel consumption. However, as it is deeply discussed in Baltagi et al. (2003) in a similar study, the use of heterogeneous models leads to highly variable and unstable estimates. Therefore, we employ a similar estimation procedure as the one by Li et al. (2014).

Table 4.1: Standard deviation of the main variables (I = 50, T = Jan’11-Oct’14)

<table>
<thead>
<tr>
<th>Pooled sample</th>
<th>log(cons.)</th>
<th>log(pre-tax price)</th>
<th>log(tax)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.7022</td>
<td>0.0486</td>
<td>0.2650</td>
</tr>
<tr>
<td>Diesel A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between</td>
<td>0.6992</td>
<td>0.0216</td>
<td>0.2507</td>
</tr>
<tr>
<td>Within</td>
<td>0.1175</td>
<td>0.0437</td>
<td>0.0929</td>
</tr>
<tr>
<td>Diesel B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.5786</td>
<td>0.0469</td>
<td>0.0623</td>
</tr>
<tr>
<td>Between</td>
<td>1.0242</td>
<td>0.0198</td>
<td>0.0136</td>
</tr>
<tr>
<td>Within</td>
<td>0.2783</td>
<td>0.0426</td>
<td>0.0608</td>
</tr>
<tr>
<td>Unleaded 95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>0.8672</td>
<td>0.0599</td>
<td>0.2088</td>
</tr>
<tr>
<td>Between</td>
<td>0.8626</td>
<td>0.0274</td>
<td>0.2043</td>
</tr>
<tr>
<td>Within</td>
<td>0.1501</td>
<td>0.0534</td>
<td>0.0517</td>
</tr>
<tr>
<td>Unleaded 98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>1.0672</td>
<td>0.0563</td>
<td>0.2218</td>
</tr>
<tr>
<td>Between</td>
<td>1.0318</td>
<td>0.0278</td>
<td>0.2181</td>
</tr>
<tr>
<td>Within</td>
<td>0.3086</td>
<td>0.0491</td>
<td>0.0503</td>
</tr>
</tbody>
</table>

\(^{29}\)Summary statistics for all the variables are included in the Appendix (Table C.1).
4.3 Results and discussion

4.3.1 The baseline model

Tables 4.2 and 4.3 present the results of the estimation of equation 4.1 for the four types of fuel considered. For all of them we estimate the model without control variables, the model controlling for registered vehicles, and the model with the full set of control variables. All the specifications include province fixed effects, month fixed effects, and a test for equal effect of “pre-tax” price and tax on fuel consumption.\textsuperscript{30} Robust standard errors clustered by month are included in parentheses. We cluster them by month because we expect that errors are correlated across provinces at a given month.\textsuperscript{31}

Table 4.2 captures the effect of changes in “pre-tax” prices and taxes on unleaded gasoline consumption. In particular, for unleaded 95 gasoline –columns (1), (2) and (3)–, the “pre-tax” price-elasticity of demand is negative but not significant for the three specifications of the model. On the other hand, the elasticity of the demand for unleaded 95 gasoline due to changes in taxes is negative and significant at the 1% level in all the model specifications: -1.56 in the model with no controls; -1.39 when controlling for vehicle fleet; and -1.33 when including all the control variables.

The results are equally striking for unleaded 98 gasoline, which are included in columns (4), (5) and (6). Indeed, the tax-elasticity of demand is negative and significant at the 1% level in all the specifications: -3.90 with no controls, -3.50 when controlling for vehicles and -3.42 when we include all the control variables. The tax-exclusive price elasticity of demand is negative but not significant for the three specifications of the model.

\textsuperscript{30}I.e. we test $\beta_1 = \beta_2$ in all the specifications. The validity of this test relies on the fact that, in Spain, around 50% of the gasoline (tax-inclusive) prices are just taxes (see the summary statistics in the Appendix, Table C.1). Therefore, we can compare the response of log prices (before tax) with log taxes.

\textsuperscript{31}We expect that fuel demand (consumption) and fuel prices are likely to be affected by common shocks at a given month in all provinces.
Table 4.2: Effects of prices and taxes on the consumption of unleaded fuels

<table>
<thead>
<tr>
<th></th>
<th>(log)Unleaded 95 cons.</th>
<th>(log)Unleaded 98 cons.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(log)Unleaded 95 pre-tax price</td>
<td>-0.0099 (0.1333)</td>
<td>-0.0636 (0.1254)</td>
</tr>
<tr>
<td>(log)Unleaded 95 tax</td>
<td>-1.5617*** (0.0864)</td>
<td>-1.3860*** (0.0860)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.3299*** (0.1034)</td>
</tr>
<tr>
<td>(log)Unleaded 98 pre-tax price</td>
<td>-0.4377 (0.3498)</td>
<td>-0.4938 (0.3503)</td>
</tr>
<tr>
<td>(log)Unleaded 98 tax</td>
<td>-3.8919*** (0.2227)</td>
<td>-3.5041*** (0.1763)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-3.4227*** (0.1676)</td>
</tr>
<tr>
<td>(log)Unleaded Vans &amp; Trucks</td>
<td>0.0021 (0.0014)</td>
<td>0.0092+ (0.0057)</td>
</tr>
<tr>
<td></td>
<td>0.0023 (0.0015)</td>
<td>0.0092+ (0.0058)</td>
</tr>
<tr>
<td>(log)Unleaded Buses</td>
<td>0.0254*** (0.0052)</td>
<td>0.0661*** (0.0156)</td>
</tr>
<tr>
<td></td>
<td>0.0264*** (0.0053)</td>
<td>0.0665*** (0.0151)</td>
</tr>
<tr>
<td>(log)Unleaded Cars</td>
<td>0.0013 (0.0024)</td>
<td>0.0086 (0.0074)</td>
</tr>
<tr>
<td></td>
<td>0.0012 (0.0023)</td>
<td>0.0085 (0.0073)</td>
</tr>
<tr>
<td>(log)Unleaded Motorcycles</td>
<td>0.0021 (0.0019)</td>
<td>0.0037 (0.0046)</td>
</tr>
<tr>
<td></td>
<td>0.0018 (0.0019)</td>
<td>0.0034 (0.0047)</td>
</tr>
<tr>
<td>(log)Unleaded Tractors</td>
<td>0.0974*** (0.0117)</td>
<td>0.2368*** (0.0641)</td>
</tr>
<tr>
<td></td>
<td>0.0916*** (0.0130)</td>
<td>0.2259*** (0.0605)</td>
</tr>
<tr>
<td>(log)Unleaded Other Vehicles</td>
<td>0.0054 (0.0035)</td>
<td>0.0100* (0.0053)</td>
</tr>
<tr>
<td></td>
<td>0.0052 (0.0034)</td>
<td>0.0097* (0.0052)</td>
</tr>
<tr>
<td>(log)Credit</td>
<td>0.0191** (0.0075)</td>
<td>0.0108 (0.0133)</td>
</tr>
<tr>
<td>(log)Housing</td>
<td>0.0258 (0.0170)</td>
<td>0.0424 (0.0298)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time FE</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Province FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>p-value ($\beta_1 = \beta_2$)</th>
<th>0.0000</th>
<th>0.0000</th>
<th>0.0001</th>
<th>0.0000</th>
<th>0.0000</th>
<th>0.0000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>2300</td>
<td>2161</td>
<td>2161</td>
<td>2300</td>
<td>2161</td>
<td>2161</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.991</td>
<td>0.991</td>
<td>0.991</td>
<td>0.971</td>
<td>0.972</td>
<td>0.972</td>
</tr>
</tbody>
</table>

Standard errors (in parentheses) are clustered by month

$+ p < 0.15, * p < 0.10, ** p < 0.05, *** p < 0.01$

The differences in the parameters that we find as well as the rejection of the proposed test in all the specifications of the model (p-values are equal to or very close to 0) suggest us that the effect of tax changes in unleaded gasoline demand is greater than the effect of price changes in this fuel demand.

Next, Table 4.3 reports the results for diesel fuel consumption. First of all, notice
that for the case of diesel B (agricultural) –columns (4), (5) and (6)– we control just for tractors, but not for the other types of vehicles. On the other hand, for the diesel A (regular) case –columns (1), (2) and (3)–, we include all types of vehicles as controls, but not tractors. The reason is that in Spain agricultural diesel (diesel B) has a lower tax rate than regular diesel (diesel A). In other words, diesel B has a tax advantage in comparison to diesel A. However, in order to buy diesel B it is necessary to prove that the purchased fuel is going to be used to power a tractor, being prohibited for other vehicles. Following this mandate, we expect that all car, motorcycle and small truck and van owners would not buy diesel B (because is prohibited), while all tractor owners would buy diesel B (because is cheaper).

The results are different from those in the previous table. First, regarding diesel A (regular) consumption –columns (1), (2) and (3)– we find a substantial difference between the “pre-tax” elasticity and the tax elasticity of the demand in the model with no controls. In fact, for this case, the “equal-effect” test is rejected at the 5% level. However, when we incorporate the control variables, this difference vanishes and the asymmetric responses are no longer observed. Thus, in the model that incorporates the full set of control variables, the “pre-tax” price-elasticity of demand is -0.33, while the elasticity of the demand due to changes in taxes is -0.48. In this case, the test for equal coefficients cannot be rejected, so we accept the alternative hypothesis that both changes in prices and taxes have a similar negative impact on the demand.

However, for the diesel B case–columns (4), (5) and (6)– we find similar evidence as for the unleaded gasoline. The coefficients for the tax-exclusive price are negative but not significant, whereas the coefficients for the tax are significant at the 1% level (around -2.4 for the three specifications of the model). Again, the “equal-effect” tests are rejected at the 1% in columns (4) and (5), and rejected at the 5% level in column (6) (with the full set of control variables).

32This is included in the Spanish Ley 38/1992, de Impuestos Especiales.
Table 4.3: Effects of prices and taxes on the consumption of diesel fuels

<table>
<thead>
<tr>
<th></th>
<th>(log)Diesel A cons.</th>
<th>(log)Diesel B cons.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) (2) (3) (4) (5) (6)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel A pre-tax price</td>
<td>-0.2980*** -0.3373*** -0.3279***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0934) (0.0998) (0.1034)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel A tax</td>
<td>-0.5752*** -0.5301*** -0.4796***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0400) (0.0359) (0.0336)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel B pre-tax price</td>
<td>-0.4174 -0.5004 -0.5137</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.3507) (0.3655) (0.3921)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel B tax</td>
<td>-2.4538*** -2.4213*** -2.3940***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.2349) (0.2484) (0.3600)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel Vans &amp; Trucks</td>
<td>0.0004 0.0002</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0016) (0.0017)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel Buses</td>
<td>0.0084** 0.0064+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0036) (0.0040)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel Cars</td>
<td>0.0048* 0.0041+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0025) (0.0024)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel Motorcycles</td>
<td>-0.1433*** -0.1135***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0263) (0.0251)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel Other Vehicles</td>
<td>-0.0061+ -0.0062*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0036) (0.0034)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel Tractors</td>
<td>0.0039 0.0037</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0049) (0.0049)</td>
<td></td>
</tr>
<tr>
<td>(log)Credit</td>
<td>0.0237***          0.0054</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0041) (0.0099)</td>
<td></td>
</tr>
<tr>
<td>(log)Housing</td>
<td>0.0732*** -0.0080</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0167) (0.0144)</td>
<td></td>
</tr>
<tr>
<td>(log)Agric. Unemployment</td>
<td>-0.0139          (0.0667)</td>
<td></td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes Yes Yes Yes Yes Yes</td>
<td></td>
</tr>
<tr>
<td>Province FE</td>
<td>Yes Yes Yes Yes Yes Yes</td>
<td></td>
</tr>
<tr>
<td>p-value ($\beta_1 = \beta_2$)</td>
<td>0.0196 0.1035 0.2093 0.0022 0.0050 0.0187</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2300 2250 2250 2208 2160 2160</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.983 0.983 0.984 0.880 0.881 0.881</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors (in parentheses) are clustered by month

* $p < 0.15$, ** $p < 0.10$, *** $p < 0.05$, **** $p < 0.01$

We end this subsection with some concluding remarks about these results. First, for the most commonly used fuels for transportation purposes (i.e. unleaded regular
gasoline and regular diesel fuel), income-related variables are relevant. In particular, for the case of the unleaded regular gasoline, the credit variable is positive and significant at the 5% level. For the regular diesel case, both the housing and the credit variables are positive and significant at the 1% level. This implies that, as one could expect, an increase in income (and their proxies) pushes up consumption of regular fuels used for individual’s transportation purposes. On the other hand, credit and housing are not statistically different from zero for the premium unleaded gasoline. Typically, premium gasoline is used to run luxury cars. Therefore, it makes sense that for luxury car owners, income variables play a marginal role on gasoline consumption decisions. Moreover, credit and housing are also not significant for the agricultural diesel case. This is in line with the fact that its consumption is linked to agricultural production (not to income).33

Second, our results for both unleaded 95 gasoline and diesel A, which are the two main fuels used for transportation purposes in Spain, are similar to those in the paper by Li et al. (2014). On the other hand, the estimated impact of the tax on consumption of unleaded 98 gasoline and diesel B are much greater. Thus, with our analysis, we are able to identify that those consumers using premium gasoline as well as the usage of gasoline linked purely to professional motives (tractors) suffer a much greater impact due to changes in prices than the fuel types used by the average car, motorcycle or small truck driver.

Finally, and following the suggestions by Hsing (1990) (and as Hughes et al. (2008) do), we have checked that these conclusions are identical when we use a log-linear regression model.

33In fact, for the diesel B case, none of the additional controls included in column (6) are significant. This fact justifies the robustness check we propose in subsection 4.2.3.
### 4.3.2 Partial adjustment models in gasoline demand

Tables 4.4 and 4.5 provide the results of the partial adjustment model specified in equation 4.5 for both unleaded gasoline fuels and diesel fuels respectively. We include the same control variables that we included in the baseline case and, again, robust standard errors clustered by month are in parentheses.

<table>
<thead>
<tr>
<th>Table 4.4: Partial adjustment model for unleaded fuels</th>
</tr>
</thead>
<tbody>
<tr>
<td>(log)Unleaded 95 cons. (t-1)</td>
</tr>
<tr>
<td>(log)Unleaded 95 pre-tax price</td>
</tr>
<tr>
<td>(log)Unleaded 95 tax</td>
</tr>
<tr>
<td>(log)Unleaded 98 cons. (t-1)</td>
</tr>
<tr>
<td>(log)Unleaded 98 pre-tax price</td>
</tr>
<tr>
<td>(log)Unleaded 98 tax</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Full set of controls</th>
<th>No</th>
<th>Yes</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Province FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2250</td>
<td>2113</td>
<td>2250</td>
<td>2113</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.994</td>
<td>0.994</td>
<td>0.974</td>
<td>0.975</td>
</tr>
</tbody>
</table>

First, Table 4.4 includes the results for the unleaded gasoline fuels. Again, the effect of a change in the tax on unleaded 95 gasoline demand is negative and significant at the 1% level if no controls are included –column (1)– and with the full set of controls –column (2). The coefficients obtained are -0.86 and -0.75 respectively. However, the coefficients for the tax-exclusive price variable are negative but not significant in both cases. The results are very similar for the unleaded 98 case –columns (3) and (4). In
fact, we obtain that tax coefficients are negative and significant at the 1% level, but “pre-tax” coefficients are negative and not statistically significant.

Second, Table 4.5 provides the results of the new model for diesel fuels. For the diesel B case—columns (3) and (4)—the results are similar to the baseline model. Both tax and “pre-tax” price coefficients are negative, but the coefficients of the tax-exclusive price are not significant, while the coefficients of the taxes are statistically significant at the 1% level.

However, the results for the diesel A case are again different. In fact, when we include all the control variables, the “pre-tax” price-elasticity of demand is -0.20 and is statistically significant at the 1% level, while the elasticity of the demand due to changes in taxes is more negative (-0.31) and significant at the 15% level. Therefore, again, the empirical evidence suggests us that there are no asymmetric reactions in regular diesel consumption due to changes in tax-exclusive prices and taxes.

The use of the partial adjustment model is not standard across the literature. However, for our particular study, the partial adjustment model seems to capture much better the consumption of diesel A, since the consumption of diesel A is more persistent and thus its consumption depends highly on the previous period consumption. In other words, considering that diesel A is cheaper than unleaded gasoline, frequent drivers tend to purchase diesel cars rather than unleaded gasoline cars. Therefore, regarding “diesel drivers” (frequent drivers), we expect that the fuel consumption in the previous period plays a major role. Thus, the introduction of the lagged consumption of diesel A mitigates the price and tax effect differences. This result is consistent with the difference in the unleaded gasoline tax revenue and diesel tax revenue found by Verboven (2002).

\[ ^{34}\text{More precisely, it is even statistically significant at the 11\% level.} \]

\[ ^{35}\text{More technically, Verboven (2002) points out that “T]he diesel engine has a higher ‘quality’ in the sense that it consumes less fuel per mile and requires less expensive fuel”}. \]
### Table 4.5: Partial adjustment model for diesel fuels

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(log)Diesel A cons.</td>
<td>0.7357***</td>
<td>0.7147***</td>
<td>0.7290*</td>
<td>0.2490</td>
<td>-0.2321***</td>
<td>-0.1996***</td>
<td>-0.4785</td>
<td>-1.2469***</td>
</tr>
<tr>
<td></td>
<td>(0.0785)</td>
<td>(0.0778)</td>
<td>(0.1717)</td>
<td>(0.0521)</td>
<td>(0.0521)</td>
<td>(0.0489)</td>
<td>(0.5212)</td>
<td>(0.2330)</td>
</tr>
<tr>
<td>(log)Diesel A pre-tax price</td>
<td>-0.2490</td>
<td>-0.3061+</td>
<td>-0.2490</td>
<td>-0.4785</td>
<td>-0.2321***</td>
<td>-0.1996***</td>
<td>-0.4785</td>
<td>-1.2469***</td>
</tr>
<tr>
<td></td>
<td>(0.1717)</td>
<td>(0.1728)</td>
<td>(0.1717)</td>
<td>(0.0521)</td>
<td>(0.0521)</td>
<td>(0.0489)</td>
<td>(0.5212)</td>
<td>(0.2330)</td>
</tr>
<tr>
<td>(log)Diesel B cons.</td>
<td>0.5259***</td>
<td>0.5241***</td>
<td>0.5259**</td>
<td>0.5259**</td>
<td>0.5259***</td>
<td>0.5259***</td>
<td>0.5259**</td>
<td>0.5259***</td>
</tr>
<tr>
<td></td>
<td>(0.0622)</td>
<td>(0.0633)</td>
<td>(0.0622)</td>
<td>(0.0622)</td>
<td>(0.0622)</td>
<td>(0.0622)</td>
<td>(0.0622)</td>
<td>(0.0622)</td>
</tr>
<tr>
<td>(log)Diesel B pre-tax price</td>
<td>-0.4785</td>
<td>-0.5516</td>
<td>-0.4785</td>
<td>-0.5516</td>
<td>-0.4785**</td>
<td>-0.5516</td>
<td>-0.5516</td>
<td>-0.5516</td>
</tr>
<tr>
<td></td>
<td>(0.5212)</td>
<td>(0.5475)</td>
<td>(0.5212)</td>
<td>(0.5475)</td>
<td>(0.5212)</td>
<td>(0.5475)</td>
<td>(0.5212)</td>
<td>(0.5475)</td>
</tr>
<tr>
<td>(log)Diesel B tax</td>
<td>-1.2469***</td>
<td>-1.2525***</td>
<td>-1.2469</td>
<td>-1.2525**</td>
<td>-1.2469***</td>
<td>-1.2525***</td>
<td>-1.2525**</td>
<td>-1.2525***</td>
</tr>
<tr>
<td></td>
<td>(0.2330)</td>
<td>(0.2520)</td>
<td>(0.2330)</td>
<td>(0.2520)</td>
<td>(0.2330)</td>
<td>(0.2520)</td>
<td>(0.2520)</td>
<td>(0.2520)</td>
</tr>
<tr>
<td>Full set of controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Province FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2250</td>
<td>2200</td>
<td>2160</td>
<td>2112</td>
<td>2250</td>
<td>2200</td>
<td>2160</td>
<td>2112</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.992</td>
<td>0.992</td>
<td>0.916</td>
<td>0.917</td>
<td>0.992</td>
<td>0.992</td>
<td>0.916</td>
<td>0.917</td>
</tr>
</tbody>
</table>

* Agricultural unemployment was not included in the set of control variables for the Diesel B case.

Standard errors (in parentheses) are clustered by month.

$+ p < 0.15, * p < 0.10, ** p < 0.05, *** p < 0.01$

### 4.3.3 Lagged effect of unemployment on diesel B consumption

Table 4.6 provides the results of the model with the lagged agricultural unemployment variable –equation 4.6– for the diesel B case. Again, the same set of control variables is included in our analysis, namely the number of registered tractors powered by diesel fuel and the credit and housing variables. Robust standard errors clustered by month are included in parentheses.

In all the alternative specifications of the model, the coefficients for the “pre-tax” price are negative. In particular, the coefficient ranges from -0.68 when the full set of controls is included –column (3)– to -0.62 when including just the contemporaneous...
Table 4.6: Model with lagged agricultural unemployment for diesel B

<table>
<thead>
<tr>
<th></th>
<th>(log)Diesel B pre-tax price</th>
<th>(log)Diesel B tax</th>
<th>(log)Agric unemployment (t-1)</th>
<th>(log)Agric unemployment (t)</th>
<th>(log)Diesel Tractors</th>
<th>(log)Credit</th>
<th>(log)Housing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>(log)Diesel B pre-tax price</td>
<td>-0.6615</td>
<td>-0.6151</td>
<td>-0.6806</td>
<td>(0.4324)</td>
<td>(0.4426)</td>
<td>(0.4600)</td>
<td></td>
</tr>
<tr>
<td>(log)Diesel B tax</td>
<td>-2.0970***</td>
<td>-2.2114***</td>
<td>-2.2404***</td>
<td>(0.3705)</td>
<td>(0.3629)</td>
<td>(0.3617)</td>
<td></td>
</tr>
<tr>
<td>(log)Agric unemployment (t-1)</td>
<td>-0.1179+</td>
<td>-0.2510**</td>
<td>-0.2180**</td>
<td>(0.0658)</td>
<td>(0.0925)</td>
<td>(0.0891)</td>
<td></td>
</tr>
<tr>
<td>(log)Agric unemployment (t)</td>
<td>0.1738+</td>
<td>0.1522</td>
<td></td>
<td>(0.1011)</td>
<td>(0.1027)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(log)Diesel Tractors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(log)Credit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(log)Housing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Province FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2160</td>
<td>2160</td>
<td>2112</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.881</td>
<td>0.882</td>
<td>0.882</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard errors (in parentheses) are clustered by month

$^+$ p < 0.15, $^*$ p < 0.10, $^{**}$ p < 0.05, $^{***}$ p < 0.01

and the lagged agricultural unemployment variable –column (2). However, these coefficients are not statistically significant. On the other hand, the tax coefficients are negative and significant at the 1% level. In particular, the tax coefficient ranges from -2.21, when no controls are included –column (1)– to -2.24, when the full set of control variables is included –column (3).

Finally, notice that in this case, the unemployment in the previous month is negative and significant in all the specifications of the model. Indeed, we find evidence of a lagged effect of unemployment in the agricultural sector in the consumption of diesel for agricultural purposes. Again, the tractor fleet, the housing and the credit variables play a marginal role.
4.3.4 **Instrumental variable regression**

This subsection discusses the main findings of the proposed instrumental variable regression. Recall that as an instrument for fuel prices we use the price of Brent crude oil interacted with monthly CPI data at the province level.

First of all, to assess the strength of our instrument, Table 4.7 includes the coefficient estimates when we regress the log of real gasoline price on the log of crude oil interacted with monthly data on CPI per province (the first stage of the IV regression). Indeed, the instrument is significant both with and without control variables at the 1% level for all fuel types.

Next, we turn to the estimation of the second stage. For the four fuel types considered, we include the results of the regression with no control variables (odd columns) and with the full set of control variables (even columns).
Table 4.7: IV regression for unleaded and diesel fuels

<table>
<thead>
<tr>
<th></th>
<th>Unleaded 95</th>
<th>Unleaded 98</th>
<th>Diesel A</th>
<th>Diesel B</th>
</tr>
</thead>
<tbody>
<tr>
<td>(log)Fuel pre-tax price</td>
<td>0.0821*</td>
<td>-0.0016</td>
<td>-0.2823**</td>
<td>-0.3977***</td>
</tr>
<tr>
<td></td>
<td>(0.0474)</td>
<td>(0.0481)</td>
<td>(0.1136)</td>
<td>(0.11154)</td>
</tr>
<tr>
<td>(log)Fuel tax</td>
<td>-1.6004***</td>
<td>-1.3581***</td>
<td>-3.9547***</td>
<td>-3.4648***</td>
</tr>
<tr>
<td></td>
<td>(0.0396)</td>
<td>(0.0501)</td>
<td>(0.0897)</td>
<td>(0.1135)</td>
</tr>
<tr>
<td>Full set of controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Time FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Province FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>2300</td>
<td>2161</td>
<td>2300</td>
<td>2161</td>
</tr>
</tbody>
</table>

First stage IV estimates: dependent variable is pre-tax price

| (log)Brent       | 0.6412***   | 0.6876***   | 0.5840*** | 0.6289*** | 0.6434*** | 0.6559*** | 0.5847*** | 0.6013*** |
|                 | (0.0068)    | (0.0071)    | (0.0063)  | (0.0066)  | (0.0054)  | (0.0058)  | (0.0052)  | (0.0058)  |
| F-test (p-value) | 0.0000      | 0.0000      | 0.0000    | 0.0000    | 0.0000    | 0.0000    | 0.0000    | 0.0000    |

F-test is for the null hypothesis that the coefficients for the interaction terms are (jointly) equal to zero
+ p < 0.15, * p < 0.10, ** p < 0.05, *** p < 0.01
First, notice that for the unleaded gasoline 95 –columns (1) and (2)– there is no evidence of a negative impact of prices on consumption, but there is substantial evidence (at the 1% level) of a negative impact of taxes on consumption. Second, for the unleaded gasoline 98 case –columns (3) and (4)– we find that the “pre-tax” price coefficient is negative and significant, but the tax coefficient is still significant and much more negative. In particular, when no controls are included, the estimated tax-exclusive elasticity of the demand is -0.28, while the estimated tax elasticity is -3.95. When including the full set of control variables, the coefficient of the tax-exclusive price is -0.40, while the estimated tax elasticity is -3.46. These findings are consistent with Hypothesis 1.

The results for the diesel B (agricultural) case –columns (7) and (8)– are fairly similar to the unleaded gasoline 98 case. Both tax-exclusive and tax coefficients are negative and statistically significant. However, the price coefficient is approximately 11 times smaller than the tax coefficient when no controls are included, while the price coefficient is 7.5 smaller than the tax coefficient when all the control variables are included.

Finally, once again, we do not find empirical evidence to support the existence of asymmetric responses in regular diesel consumption –columns (5) and (6). As we can see, both “pre-tax” coefficients and tax coefficients are negative and significant at the 1% level, and very similar in magnitude: all of them lie in between -0.56 and -0.41. Therefore, the instrumental variable regression refutes the findings in the previous subsections.

36 Notice that we can compare both effects (in level terms) because the instrument is expressed in euro per litre and, as shown in the Appendix (Table C.1), the mean and variance of the instrument and the taxes are fairly similar in magnitude.
4.4 Conclusions and policy implications

Policy makers have previously thought that, since gasoline demand has been proven to be price inelastic—see Akinboade et al. (2008), Baranzini and Weber (2013), Brons et al. (2008), Galindo (2005), Havranek et al. (2012), Lin and Zeng (2013) and Ramanathan (1999), among others—raising taxes on gasoline can lead to a substantial increase in the tax revenue, helping to reduce the budget deficit.

In this chapter we examine the elasticity of the demand for different fuels using monthly data from 2011 until 2014 at the regional level in Spain taking into account separately changes in tax-exclusive fuel prices and changes in taxes on fuels. Contrary to what was previously thought, we find that while changes in “pre-tax” prices has little or no impact on gasoline demand, changes in taxes produce much larger and significant changes in gasoline consumption. This result holds for unleaded gasoline consumption (regular and premium) and diesel fuel used for agricultural purposes. Moreover, we find that our result is robust to alternative specifications of the main model that take into account potential dynamic adjustments in the consumption patterns and is also validated when we instrument for fuel prices.

On the other hand, this asymmetric behavior is not observed in the regular diesel demand. In fact, consumers’ responses to changes on prices and taxes of diesel are fairly similar. This fact is consistent with the previous papers that have studied the different incidence of taxes on different fuels—see Verboven (2002).

As a “take home message”, notice that these findings have several and notorious implications in terms of energy policy. First, our results cast doubt on the effectiveness of taxing at the pump as a policy instrument to increase revenue collection—especially if the taxes are imposed on unleaded gasoline consumption. Therefore, if the goal of the government is to increase the tax revenue (as is the case in most of the examples included in the introduction, including the Spanish case) we conjecture that it might be more effective to implement a different tax rather than a “tax
at the pump” –see Goulder (1994). Second, as car dealers price discriminate when selling cars by pricing differently diesel and unleaded cars, we suggest governments to “tax-discriminate” when imposing taxes at the pump. In other words, since diesel and unleaded consumers react differently to changes in taxes, it may be efficiency-enhancing to impose different rates according to their consumption responsiveness. Third, our results also cast doubts on previous studies that estimated the effect of taxes based on the “overall” (i.e. the tax-inclusive) elasticity of the gasoline demand. By using the “overall” elasticities instead of “separated” ones, policymakers overestimate the impact on revenue of an increase in the tax at the pump and underestimate the impact on pollution and climate change –see Li et al. (2014).
Chapter 5

Where to locate a gas station?
Socially optimal location in a linear city game with non-uniform distribution of consumers

5.1 Introduction

One of the most well-known models of spatial competition in the industrial organization literature is the linear city model, introduced by Hotelling (1929) and refined by d’Aspremont et al. (1979). For decades, this model has received the attention of many authors, especially the theoretical ones. In this chapter we want to demonstrate that there is also room for some empirical work using Hotelling’s model. In particular, building upon recent literature, we explore price competition in a linear city model to see what we can learn about the distribution of consumers and consumers’ transportation cost—or “travel cost”. We develop a strategy that, with sufficient variation in the data, allows us to recover points in the Cartesian plane of the distri-
duction of consumers. Using such points we get a very precise approximation of the
ture distribution and the transportation cost parameter. Once we obtain both the
distribution of consumers and the transportation cost, we are able to assess what is
the socially optimal location of a potential entrant. In addition, we also propose some
counterfactual exercises.

Perhaps the most implausible assumption in the original setting by Hotelling
(1929) is the uniform distribution of consumers along the unit interval. In fact,
since its original formulation, many authors have deviated from this assumption and
have characterized the market equilibrium assuming instead that consumers are non-
uniformly distributed—for instance, Neven (1986), Tabuchi and Thisse (1995), Anderson et al. (1997), Shuai (2014). As these (and many other) authors noticed, consumers
usually concentrate on certain sections of a straight interval—for instance, around the
center of the interval. Following the trend in the literature, we depart from the origi-
nal model by Hotelling (1929) by allowing a “more realistic” non-uniform distribution
of consumers along the straight interval.

Assuming that the distribution of consumers falls in the general category of log-
concave functions, we describe the best response functions in the Hotelling’s game
with exogenous locations. Using the unique equilibrium conditions, we develop our
identification strategy to approximate the distribution of consumers. Furthermore, if
we restrict such a distribution to fall in the class of beta distributions, we are also
able to obtain the transportation cost (denoted \( \tau \)). For that purpose, we estimate
the shape parameters of the beta distribution via nonlinear least squares. Then, we
estimate \( \tau \) by minimizing the sum of squared distance (errors) between the estimated
beta distribution and the identified points.

As an application, we use a dataset of daily prices and costs of a pair of gas

\(^{1}\)Needless to say, our technique does not perform as well under extreme conditions. For instance, if all consumers are located on a point in the interval.

\(^{2}\)As we explain later, most of the commonly used distributions satisfy the log-concavity assumption, including the uniform distribution, the beta distribution, the normal distribution, etc.
stations located in a straight highway in Spain. The variation in the data allows us to recover the distribution of consumers between two points (two towns) along a straight highway.\textsuperscript{3} Thus, we are able to indicate in which exact point in the highway the government should give permission to build a new gas station to maximize consumers’ welfare.

The Spanish retail gasoline market is an appropriate market to study the proposed model for several reasons. To begin with, the geography of Spain makes this country an ideal place to evaluate the Hotelling framework. In fact, the Spanish topography consists of a wide and broad plateau in the center of the country of over 154,440 square miles. Such a plateau occupies around 50% of the Iberian Peninsula. This topography favors the construction of straight roads, which are appropriate scenarios to test the linear city model.

Second, gasoline has usually been considered a homogeneous product –see Tappata (2009), Wang (2009), Chandra and Tappata (2011), Janssen et al. (2011).\textsuperscript{4} Therefore, we leave aside the challenges generated by product differentiation –other than just gas station location– or unobserved product characteristics. Moreover, the Spanish retail gasoline market is a highly concentrated one. In fact, five major operators controlled about 75-80% of the market in 2008 –see Bello and Contín-Pilart (2012). Therefore, it is relatively easy to find highly-recognized brands controlling pair of gas stations in highways, ruling out differences due to asymmetric quality perceptions across brands –see Lewis (2008).

Finally, in many highways in Spain there is a lack of gas stations. For instance,

\textsuperscript{3}We carefully selected a straight section of the highway with enough entries and exits along it, which generate different demand for gas in different points. In addition, we consider a highway for which there are no obvious or equally convenient alternatives routes.

\textsuperscript{4}The fact that gasoline is a homogeneous good is challenged by Lewis (2008). However, his argument relies on the price differences that arise between high-brand and low-brand retailers in highly dense urban areas considering how many and of which type are the surrounding competitors. To avoid the existence of differences between competitors, we focus on pair of gas stations with similar characteristics (namely, highly-recognized brands, same opening hours, etc.) and in long and straight sections of highways served only by a couple of gasoline retailers.
there are 186 miles with no gas stations in the A-6 Highway, also called the “Northwestern Highway” (*Autovía del Noroeste*), one of the main highways in Spain that connects Madrid with the Northern region of Galicia. Another case in point: there is an approximately 57-miles section in the AP-7 Toll Highway (the so-called “Mediterranean Highway” or *Autopista del Mediterráneo*) between Xeraco and Sagunto (both in the province of Valencia) without gas stations. We can find another example in the AP-66 Freeway (also called *Ruta de la Plata*) which connects León with Asturias. Even though this freeway served about 7,506 vehicles per day in 2014 according to the Spanish Ministry of Public Works and Transport\(^5\), it has just a couple of gas stations along it. In particular, if you travel northbound, there is a 38-miles section with no gas stations. Likewise, if you travel southbound, there are about 34 miles with no gas stations.\(^6\)

It is often argued that the lack of gas stations in Spanish roads is due to excessive bureaucracy. In fact, to open a new gas station in a Spanish highway, it is mandatory to get an authorization issued and approved by the General Direction of Highways (*Dirección General de Carreteras*), which is a body of the Ministry of Public Works and Transport. The General Direction of Highways studies both the potential location along the highway and the viability of the project (among many other things), and decides whether a new gas station should or should not be opened in such location. This decision is legally binding.\(^7\) Therefore, using our methodology, we add some light to the decision process of the location in which a new gas station should be opened based on consumers’ welfare.


\(^6\)Some other examples: the AP-52, also called “Central Galician Freeway” (*Autopista central gallega*), which is 35.2 miles long, has just one gas station (close to the town of Sillero) serving the 5,662 vehicles that circulate on average through it. Another 44-miles section of the aforementioned AP-7 in the province of Murcia has no gas stations.

\(^7\)Further details regarding this requirement can be found in the Spanish legislation; more precisely in the General Regulation of Highways Act (*Real Decreto 1812/1994*), passed on September 2, 1994.
linear city model of spatial competition in prices. Several previous authors have empirically validated the linear city model—for instance, Liarte and Forgues (2008) and Torrisi (2008). However, the methodology we propose goes beyond simply validation, by allowing us to recover the (potentially non-uniform) distribution of consumers. In fact, this chapter contributes to the IO literature by introducing a relatively simple identification approach using a linear city game framework.

The present chapter combines ideas in previous literature. In particular, we have similar best-response functions as the ones in Anderson et al. (1997) and Shuai (2014). As they do, we also consider that the distribution of consumers is log-concave. This assumption, which is satisfied by most of the commonly used distribution functions and that is weaker than imposing concavity, guarantees uniqueness of equilibrium in the static game, as shown by Caplin and Nalebuff (1991).

The rest of the chapter is organized as follows. In the next subsection we briefly discuss some related studies. In section 5.2, we introduce the model of price competition in a linear city model. Section 5.3 discusses the challenges we face in our model application. Section 5.4 employs our methodology to study price competition in pair of gas stations in straight highways in Spain. We recover the distribution of consumers that purchase gas from these gas stations in subsection 5.4.3. Section 5.5 concludes.

### 5.1.1 Related literature

As we mention above, previous authors have empirically validated the linear city model. For instance, Liarte and Forgues (2008) study location choices of hamburger restaurants in Paris between 1984 and 2004. They find that while market leaders look for spatial differentiation, challengers look for agglomeration due to positive externalities of locating closer to a leader. In our chapter we remain silent about incumbents’ location choices by taking the location as exogenously given. In our case, this as-
umption is in line with the fact that the General Direction of Highways has the final say when it comes to the location decision of a gas station along a highway. Torrisi (2008) studies prices and location decisions of drink kiosks along main avenues in Catania (Italy). He assumes that consumers are distributed according to a triangular distribution. His goal is just to validate Hotelling’s theoretical results. However, neither this chapter nor the former one recover the distribution of consumers, which is our main task.

A more recent study that is closer to ours is by Moul (2015). Building on the Salop’s circular city model, he is able to identify demand parameters and firms’ variable costs. As in the original Salop model, he assumes that consumers are uniformly distributed around a unit circumference with three symmetrically located firms. As Moul (2015) does, we assume that the location of firms is exogenously given. However, we do not impose symmetric location of firms, which is a particularly restrictive assumption.

The use of a linear city model to study prices and product differentiation is in Bresnahan (1981). In particular, the author estimates sellers’ markups using U.S. automobile market data. He assumes that product differentiation is based on a single scalar, which represents the “quality of the car”. This parameter is estimated from cars’ length, weight, engine horsepower, fuel efficiency and body-type. Using a similar data set, Feenstra and Levinsohn (1995) extend Bresnahan’s markup estimation to the multi-dimensional product differentiation case. Unlike them, and considering that our task is not to estimate markups but to recover the distribution of consumers, we remain stick to the one-dimensional product differentiation case (location).

Houde (2012) also uses data on gasoline retail markets to estimate a model of spatially differentiated goods. He estimates the parameters of the model using data on fuel prices, fuel consumption, commuting paths, residence of consumers, etc. However, while Houde (2012) assumes a complex firms’ network structure at the cost of having
a very detailed dataset for many different variables, our study assumes a simpler firms’ network, which requires having less detailed data.

In a dynamic setting, Doganoglu (2003) and Laussel et al. (2004) employ Hotelling’s model to study markets with network effects. In particular, the latter paper (which is an extension of the former) studies the evolution of access prices and firms’ market shares for goods with consumption externalities – i.e. congestion effects. Unlike us, they assume that consumers are uniformly distributed along a straight interval to analyze the impact of market shares on access prices. Contrary to them, we are interested in analyzing how changes in variable costs and equilibrium prices from one period to another can lead us to learn the (potentially non-uniform) distribution of consumers.

Bajari and Benkard (2005) identify and estimate the demand of vertically differentiated products. However, the set of assumptions they use differs from ours in numerous ways. For instance, while they focus on vertically differentiated markets, we study a market with horizontal product differentiation. A more in-depth discussion of studies that estimate demand under many different scenarios and assumptions is in Ackerberg et al. (2007).

5.2 Game theoretic framework

We assume that there are two firms \((i \in \{1, 2\})\) located along the interval \([x, \overline{x}]\).\(^8\) We denote firm \(i\)’s position as \(x_i\) and we assume that they are exogenously given and constant over time. Without loss of generality we assume that \(x_1 \leq x_2\). Both firms produce and sell a homogeneous good. There is a strictly stationary mass one continuum of consumers distributed over the aforementioned interval according to some density \(f(\cdot)\) for all \(x \in [x, \overline{x}]\). We denote as \(F(\cdot)\) the cumulative distribution

\(^8\)For practical purposes, we normalize this interval to \([0, 1]\) in upcoming sections. However, we keep the general form in this section to point out that the existence and uniqueness of an equilibrium does not require such normalization.
function associated to such density function; we also assume that $F^{-1}(\cdot)$ exists. To guarantee uniqueness of the equilibrium price, we impose the following assumption.

**Assumption 1.** $f(x)$ is log-concave on the interval $[\underline{x}, \overline{x}]$. That is, $\log f(x)$ is concave.

Notice that most of the common distributions satisfy the log-concavity assumption, which is less restrictive than concavity. It also implies that both $F(\cdot)$ and $[1 - F(\cdot)]$ are log-concave, which is a direct consequence of the Prékopa-Borell Theorem.\(^9\)

Both firms and consumers make decisions at times $t = 1, 2, \ldots, \infty$. At time $t$, each firm faces a commonly observed constant marginal cost $c(\theta^t)$, which is depends on a (common) random shock, $\theta \in \Theta$. This shock is realized at the beginning of each period. That is, $c : \Theta \to \mathcal{C}$, where $\mathcal{C} \subseteq \mathbb{R}_+$ is the finite set of marginal costs.\(^10\) Given $c(\theta^t)$, firms simultaneously choose the price they charge to consumers. That is, at each period $t$ both firms post the price they charge, $p^t_i \in \mathcal{P} \subseteq \mathbb{R}_+$.

Likewise, every period $t$, consumers make consumption choices. In particular we assume that each consumer buys one unit of the homogeneous good and pays the price charged by the firm from which she buys. Moreover, as it is standard in the Hotelling’s linear city framework, we assume that consumers incur a quadratic “travel cost” ($\tau$). Therefore, the total cost for a consumer located at $x \in [\underline{x}, \overline{x}]$ of buying the homogeneous product from firm $i$ at time $t$ is given by $p^t_i + \tau(x - x_i)^2$.

A consumer located at $\tilde{x}^t \in [\underline{x}, \overline{x}]$ is indifferent between buying the homogeneous good from firm 1 and buying the homogeneous good from firm 2 at time $t$ if $p^t_1 + \tau(\tilde{x}^t - x_1)^2 = p^t_2 + \tau(\tilde{x}^t - x_2)^2$. Thus, at each $t$, the indifferent consumer is uniquely determined as follows:

\(^9\)See Caplin and Nalebuff (1991) for a more in-depth discussion on this Theorem.

\(^10\)In our particular context, the common marginal cost for two gas stations is the rack (wholesale) price of gasoline.
\[ \tilde{x}^t(p_1^t, p_2^t; x_1, x_2) = \frac{p_2^t - p_1^t}{2\tau(x_2 - x_1)} + \frac{x_2 + x_1}{2}; \quad (5.1) \]
i.e., in period \( t \), firm 1 serves to all consumers located to the left of the indifferent consumer and firm 2 serves to all consumers located to the right of \( \tilde{x}^t(\cdot; \cdot) \).

We denote firm \( i \)'s profit as \( \pi^t_i(p_1^t, p_{-i}^t, \theta^t) \), which is characterized by the following expression:

\[ \pi^t_i(p_1^t, p_{-i}^t, \theta^t) = \int_{\tilde{x}^-}^{\tilde{x}^+} [p_1^t - c^t(\theta^t)] f(x) dx \quad (5.2) \]

where \( \tilde{x}^- = x \) and \( \tilde{x}^+ = \tilde{x}(\cdot; \cdot) \) if \( i = 1 \) and \( \tilde{x}^- = \tilde{x}(\cdot; \cdot) \) and \( \tilde{x}^+ = x \) if \( i = 2 \).

We assume that the variable cost evolves according to some low of motion which depends on the variable cost in the previous period and a random shock. That is, \( \theta^{t+1}(\theta^t, \epsilon^{t+1}) \). We explicitly assume that the current decision variable (today’s price) does not impact the evolution of the state variable (future firm’s variable cost) or conversely, that the variable cost does not depend on previous period firms’ prices choices.\(^{11}\) By doing so, the actual functional form of \( c(\theta^{t+1}) \)–which is actually unknown– loses relevance, as we demonstrate below.

Since each period both firms face an identical problem, we solve for the profit maximizing problem for both firms at every period \( t = 0, 1, 2 \cdots \). Thus, at an interior solution, firm \( i \)'s best response (i.e. \( p_i^{t\ast} > 0 \)) solves the following expression:

\[ \left[ \frac{\partial \tilde{x}^+}{\partial p_i^t} f(\tilde{x}^+) - \frac{\partial \tilde{x}^-}{\partial p_i^t} f(\tilde{x}^-) \right] (p_i^t - c(\theta^t)) + [F(\tilde{x}^+) - F(\tilde{x}^-)] = 0 \quad \forall t \quad (5.3) \]

Since we are considering the two-firm case, we can characterize equation 5.3 for

\(^{11}\)On the contrary, the assumption that the current firm’s price has an impact on the evolution of the variable cost may be interpreted in two ways in the context of gasoline retail markets. First, by doing so, we may be assuming that single gas stations’ posted price has an impact on crude oil prices. Second, we may have in mind some kind of inertia in posted prices, potentially explained by some kind of rigidity (such as menu costs). While the first explanation seems implausible (no single gas station seems to have enough power to determine the evolution of the price of fuel commodities) the second explanation lacks also of realism if we take into account that posted prices change every day and that they are posted electronically, so the cost of changing them is virtually zero.
both firms. By doing so, we get the following expressions:

\[ p_{1}^{t *} (\cdot) = 2\tau(x_2 - x_1) \frac{F(\tilde{x})}{f(\tilde{x}_t)} + c(\theta^t) \]  \hspace{1cm} (5.4)

and

\[ p_{2}^{t *} (\cdot) = 2\tau(x_2 - x_1) \frac{1 - F(\tilde{x})}{f(\tilde{x}_t)} + c(\theta^t) \]  \hspace{1cm} (5.5)

that must be satisfied in equilibrium at every \( t \).

**Proposition 15.** Under Assumption 1, \( p_{i}^{t *} (\cdot) \), the equilibrium price is unique for \( i \in \{1, 2\} \) for all \( t \).

The equilibrium best response functions for both firms (equations 5.4 and 5.5) depend on four key parameters: first, the shape of the consumers’ distribution along the interval; second, the location of the firms; third, the transportation cost for the consumers, denoted as \( \tau \); finally, they also depend on firms’ marginal costs. Notice that, in our setup, only the last one is allowed to change from one period to another. On the other hand, neither the distribution –due to strictly stationary– nor the locations –exogenously given– nor the transportation costs change from one period to another. Therefore, only the former element (variable cost) will give us the necessary variation to rationalize the changes from period to period observed in firms’ prices.\(^{12}\)

However, the other (constant) parameters play also a prominent role in determine the best response functions at every \( t \). In Appendix E we provide some examples of the different equilibrium prices we could observe depending on distribution shape, locations and “travel costs”.

\(^{12}\)A different discussion is about which are the observable and the non-observable variables. We deal with that in the following section.
5.2.1 Further theoretical aspects: entrant’s location and consumers’ welfare

As explained above, one of the goals of this study is to provide a tool that helps us determine where is the optimal location of a potential entrant in the linear city model. Therefore, a social planner will decide where to locate the potential entrant firm along the interval such that it maximizes consumers’ welfare. As Shuai (2014) does, we assume that maximizing consumers’ welfare is equivalent to minimizing consumers’ total transportation cost.\(^{13}\)

If the entrant firm is to be located between the two incumbent, then we need to find the consumer \(\hat{x}\) such that the sum of consumers’ transportation costs to the right and to the left of it are equal. That is

\[
\int_{x_1}^{\hat{x}} \tau(x - x_1)f(x)dx = \int_{\hat{x}}^{x_2} \tau(x - x_2)f(x)dx
\]

(5.6)

where \(\hat{x}\) is unique if \(F(\cdot)\) is strictly monotone in between \(x_1\) and \(x_2\).

**Claim 1. Unique socially optimal entrant’s location**

*If \(F(\cdot)\) is strictly monotone over the interval \([x_1, x_2]\), then \(\hat{x}\) is unique.*

Therefore, in order to determine where to locate the potential entrant such that the welfare is maximized (or such that total transportation costs are minimized), we need to compare the total transportation cost that consumers incur for each of the three relevant areas indicated in Figure 5.1. That is the total “travel cost” for consumers between \(\underline{x}\) and \(x_1\), denoted as \(T_1\); for consumers between \(x_1\) and \(x_2\), denoted as \(T_{\text{mid}}\), and for consumers between \(x_2\) and \(\overline{x}\), denoted as \(T_2\).

\(^{13}\)If we think in our example, the goal of the social planner (the General Direction of Highways) is to authorize the construction of a new gas station in such a way the total probability of vehicles running out of gas is minimized.
Figure 5.1: Total transportation costs for different sections of the linear city

Analytically, the total transportation costs that consumers incur in the three areas are given by the following expressions:

\[ T_1 = \int_{x_1}^{x} \tau (x - x_1)^2 f(x) \, dx \quad (5.7) \]

\[ T_{\text{mid}} = \int_{x_1}^{\hat{x}} \tau (x - x_1)^2 f(x) \, dx = \int_{\hat{x}}^{x_2} \tau (x - x_2)^2 f(x) \, dx \quad (5.8) \]

\[ T_2 = \int_{x_2}^{x} \tau (x - x_2)^2 f(x) \, dx \quad (5.9) \]

Having these expressions in mind, the socially-optimal decision rule on entrant’s location, denoted as \( x^* \), is given as follows:

a) If \( T_1 > \max\{T_{\text{mid}}, T_2\} \), the socially optimal location of the entrant is at \( x^* = \arg \min_{x \in [x_1, x]} f(x) \) such that \( f(x^*) > 0 \).\(^{14}\)

b) If \( T_{\text{mid}} > \max\{T_1, T_2\} \), the socially optimal location of the entrant is at \( x^* = \hat{x} \).

c) If \( T_2 > \max\{T_1, T_{\text{mid}}\} \), the socially optimal location of the entrant is at \( x^* = \arg \max_{x \in [x_2, x]} f(x) \) such that \( f(x^*) > 0 \).\(^{15}\)

In the cutting-edge cases in which there is a tie, we set the following tie-breaking rules:

\(^{14}\)In this case, if \( f(\cdot) \) is strictly positive in all its support, the socially optimal location of the entrant is at \( x^* = 0 \).

\(^{15}\)In this case, if \( f(\cdot) \) is strictly positive in all its support, the socially optimal location of the entrant is at \( x^* = 1 \).
d) (Tie-breaking rule #1) If $T_1 = T_{\text{mid}} > T_2$, the socially optimal location of the entrant is determined by the following rule: $x^* = \arg\min_{x \in [\underline{x}, \overline{x}]} f(x)$ such that $f(x^*) > 0$ with probability $\frac{1}{2}$ and $x^* = \hat{x}$ with probability $\frac{1}{2}$.

e) (Tie-breaking rule #2) If $T_{\text{mid}} = T_2 > T_1$, the socially optimal location of the entrant is determined by the following rule: $x^* = \hat{x}$ with probability $\frac{1}{2}$ and $x^* = \arg\max_{x \in [\underline{x}, \overline{x}]} f(x)$ such that $f(x^*) > 0$ with probability $\frac{1}{2}$.

f) (Tie-breaking rule #3) If $T_2 = T_1 > T_{\text{mid}}$, the socially optimal location of the entrant is determined by the following rule: $x^* = \arg\max_{x \in [\underline{x}, \overline{x}]} f(x)$ such that $f(x^*) > 0$ with probability $\frac{1}{2}$ and $x^* = \arg\max f(x)$ such that $f(x) > 0$ with probability $\frac{1}{2}$.

We use these decision rules to run the welfare analysis in section 5.4 by allowing a potential entrant to be located at the socially optimal location.

### 5.3 Model application

In the previous section we obtained a couple of expressions that characterize the best response functions for both firms at every period in the game. These expressions depend on four key elements, namely the shape of the distribution of consumers, firms’ locations, consumers’ transportation cost, and firms’ marginal costs. Among these elements, we assumed that only the last one (firms’ marginal costs) changes from one period to another, while the others are assumed to be constant over time.

A different discussion regarding these elements is whether they are observable or not observable. Guided by the data that we have available for the experiment that we propose in an upcoming section, we assume the following. First, we assume that the location of the firms, which is constant over time and exogenously given, is observable. Having in mind the retail gasoline market example, it is reasonable to assume that...
the location of two gas stations along a highway can be easily obtained with GPS technology, and that gas stations do not change location very frequently.\footnote{More realistic is to assume that there is entry and exit in the market instead. Although entry is not frequently observed either –maybe due to excessive bureaucracy and entry deterrence– we leave this question aside.}

Second, we also assume that the variable costs at every period $t$ are also observable. Considering that we are dealing with a homogeneous-good market, finding the cost of a (equally) homogeneous raw material is not difficult if we take into account that many of the commonly used raw materials (including gasoline) are publicly traded, so there is a commonly known price every day.

Finally, we assume that the two remaining key equilibrium elements –the distribution of consumers and the transportation cost– are not observable. First, we make a few comments on the distribution of consumers. As we mentioned in the introduction, it does not seem realistic to assume that consumers are uniformly distributed along the interval. In fact, many authors have already analyzed the Hotelling’s linear city model equilibrium with a “more realistic” non-uniform distribution of consumers.\footnote{See Neven (1986), Tabuchi and Thisse (1995), Anderson et al. (1997), Torrisi (2008), Shuai (2014).} Moreover, it is usually not easy, nor cheap, to keep track of consumers’ most frequent locations. In the highway example, although there is some data on average usage per day and traffic congestion, it is difficult to measure it by sections due to the existence of many entries and exits allowing access to nearby towns. Thus, although we may know the average number of people that use a highway per day, we do not observe which road sections present more traffic.

Finally, it seems also reasonable to assume that the transportation cost is not observable, since it is consumers’ private information. Therefore, the transportation cost parameter constitutes a primitive of our model that we estimate in the upcoming subsections.
5.3.1 Approximating the distribution of consumers under known transportation cost

Part of the problem we face is how to recover the (non-observable) distribution of consumers, i.e. $f(\cdot)$. For that purpose, we proceed as follows. First, we solve the equilibrium price expressions (equations 5.4 and 5.5) at $t$ for $F(\cdot)$. Using the resulting condition we rely on the existence of variation in the variable costs from one period to another that hence will generate variation in the equilibrium prices from one period to another. After $T$ periods, the changes in both variables will give us a sample of (at most) $T$ points in $\mathbb{R}^2$ that correspond to points in the Cartesian plane of the distribution function we are interested in. Thus, using the set of $T$ points in the Cartesian plane, we can recover an approximation of the actual distribution using smoothing techniques.

Reconsider equations 5.4 and 5.5, and denote firm $i$’s equilibrium price at $t$ as $p_{t,i}^*$, were $p_{t-i}^*(p_{t-i}^*, \theta_t)$ at period $t$, for $i \in \{1, 2\}$. Therefore, we have

$$p_{t,1}^* = 2\tau(x_2 - x_1)\frac{F(\bar{x}_t)}{f(x_t)} + c(\theta_t) \quad (5.10)$$

and

$$p_{t,2}^* = 2\tau(x_2 - x_1)\frac{1 - F(\bar{x}_t)}{f(x_t)} + c(\theta_t) \quad (5.11)$$

Solving these equations for $F(\cdot)$ we get:

$$f \left[ \frac{p_{2,t}^* - p_{1,t}^*}{2\tau(x_2 - x_1)} + \frac{x_2 + x_1}{2} \right] = \frac{2\tau(x_2 - x_1)}{p_{1,t}^* - c(\theta_t) + p_{2,t}^* - c(\theta_t)} \quad \forall t \quad (5.12)$$

Thus, if we assume that we know the transportation cost parameter ($\tau$), sufficient variation in costs from one period to another will give us different points of the density function.

Call $\lambda$ and $\nu$ a point in the range and in the domain of $f(\cdot)$ respectively. If we
get enough observations (and variation in costs and prices), we will have sufficiently
distinct points \((\lambda^1, \nu^1), \cdots, (\lambda^T, \nu^T)\), which conform a sample from the distribution. Using them, we can apply smoothing techniques to get a close approximation of the true underlying density function. This exercise, based on a simulation study, is included in Appendix F.

5.3.2 Parametric restrictions

In the previous subsection we assumed that we know the transportation cost \((\tau)\) to show how we are able to recover the distribution of the demand. However, as we also argued above, the assumption that \(\tau\) is known is quite unrealistic. The transportation cost parameter is consumers’ private information, and there is not an easy, nor obvious, way to empirically measure it.\(^{18}\) This is actually the second empirical challenge that we face in this section.

Parametric restriction of the distribution of consumers

Both the RHS and the LHS of equation 5.12 –which is the key expression in our strategy– depend on the parameter \(\tau\), which is unknown. Thus, given \(T\) observations of the equilibrium prices and marginal costs, we will have \(T\) points of the distribution that depend on the unknown \(\tau\), \([\lambda^1(\tau), \nu^1(\tau)], \cdots, [\lambda^T(\tau), \nu^T(\tau)]\). This implies that for each parameter \(\tau\), we can identify \(T\) different points in the Cartesian plane that correspond to the distribution of consumers.

In order to solve this problem and, more generally, to estimate the distribution of consumers with \(\tau\) unknown, we need to impose some parametric restrictions. In particular, we assume that consumers are distributed according to a beta distribution, with shape parameters \((\alpha, \beta)\). There are several reasons to use this distribution.

\(^{18}\)Previous authors have proposed different methodologies to estimate transportation cost parameters –see, for instance, Englin and Shonkwiler (1995). However, we need to think that \(\tau\) (the transportation cost) acts as an “equilibrium parameter” that help us fitting the model equilibrium into the observed data.
First, the beta distribution is log-concave as long as $\alpha, \beta \geq 1$. If so, Assumption 1 is therefore satisfied. Second, the support of the beta distribution is the interval $[0, 1]$. This interval is usually employed in the context of the linear city model. Third, the shape parameters of the beta distribution can produce a rich variety of shapes, yielding a symmetric, left-skewed or right-skewed distribution (with different high for peaks and length for valleys) by modifying $\alpha$ and $\beta$ respectively. Thus, the beta distribution allows us to considered many different scenarios, including the extreme and unlikely case in which consumers are uniformly distributed.$^{19}$

Estimation of the shape parameters of the distribution via nonlinear least squares

In this subsection we estimate the shape parameters of the distribution of consumers, which is assumed to belong to the class of beta distributions. We denote $\hat{\alpha}(\tau)$ and $\hat{\beta}(\tau)$ the shape parameters estimators that are based on the sample of identified points $[\lambda^1(\tau), \nu^1(\tau)], \ldots, [\lambda^T(\tau), \nu^T(\tau)]$. Since our sample of identified points depends on $\tau$, these shape parameters estimators, as indicated, also depend on $\tau$.

We use the identified sample to perform a nonlinear regression given $\tau$. In this case, the model we are interested in can be written as follows:

$$\nu^t(\tau) = f(\lambda^t(\tau), \alpha, \beta) + \varepsilon^t \quad (5.13)$$

where $f(\cdot, \alpha, \beta)$ is the density function of a beta distribution with shape parameters $\alpha$ and $\beta$, and $\varepsilon^t$ is the usual error term.

We can obtain an estimator of the shape parameters by performing nonlinear least squares (NLS), by minimizing the sum of squared errors $\sum_{t=0}^{T} [\nu^t(\tau) - f(\lambda^t(\tau), \alpha, \beta)]^2$. Given $\tau$, the NLS estimator of the shape parameters, denoted $\hat{\alpha}_{\text{NLS}}(\tau)$ and $\hat{\beta}_{\text{NLS}}(\tau)$, solve

$^{19}$This happens when both shape parameters, $\alpha$ and $\beta$, are equal to 1.
the following system of FOCs:

\[
-2 \sum_{t=0}^{T} \frac{\partial f(\cdot)}{\partial \alpha} \bigg|_{\hat{\alpha}_{nls}(\tau)} \left[ \nu^t(\tau) - f \left( \lambda^t(\tau), \hat{\alpha}_{nls}(\tau), \hat{\beta}_{nls}(\tau) \right) \right] = 0
\]  

\[
-2 \sum_{t=0}^{T} \frac{\partial f(\cdot)}{\partial \beta} \bigg|_{\hat{\beta}_{nls}(\tau)} \left[ \nu^t(\tau) - f \left( \lambda^t(\tau), \hat{\alpha}_{nls}(\tau), \hat{\beta}_{nls}(\tau) \right) \right] = 0
\]

which can be solved using numerical methods.

**Estimation of the transportation cost**

In the previous subsection we discuss how to estimate the shape parameters for any transportation cost \((\tau)\). The task in this subsection is to select the “appropriate” \(\tau\). In other words, we need to estimate the transportation cost in our model.

To do so we rely on the fact that if the real transportation cost \((\tau)\) and the estimated one (denoted as \(\hat{\tau}\)) are close to each other, the identified points and the beta distribution whose shape parameters are estimated via NLS will be close to each other too. Based on this idea, we build an estimator of the transportation cost based on the minimization of the squared discrepancies (errors) between identified points and the estimated beta distribution.

More formally, and keeping in mind the assumption that the true distribution of consumers follows a beta distribution with shape parameters \(\alpha\) and \(\beta\), if \(\tau\) is the true transportation cost parameter then:

\[
\nu^t(\tau) = f(x^t, \alpha, \beta) \quad \forall t
\]

were \(x^t\) is the value in the domain of the beta distribution whose image is \(\nu^t(\tau)\) under the beta distribution function.\(^{20}\)

However, we do not have \(\alpha\) and \(\beta\) but an estimator of them (via nonlinear least

\(^{20}\)If \(\tau\) is the true transportation cost parameter, then \(x^t = \lambda^t(\tau)\).
squares), denoted \( \hat{\alpha}^{\text{nls}}(\hat{\tau}) \) and \( \hat{\beta}^{\text{nls}}(\hat{\tau}) \). Therefore:

\[
\nu^t(\hat{\tau}) \approx f(x^t, \hat{\alpha}^{\text{nls}}(\hat{\tau}), \hat{\beta}^{\text{nls}}(\hat{\tau})) \quad \forall t
\]

where \( x^t \) is as defined above. Thus, we can obtain a precise estimator of the transportation cost by minimizing the sum of the (squared) distance between the LHS and the RHS of expression 5.16. Notice that the properties that apply to the shape parameter estimators (including consistency) also apply for the transportation cost estimator.

To facilitate the computational burden and, more importantly, to avoid erroneous solutions, we impose \( \hat{\tau} \) to belong to a set of feasible solutions, denoted as \( \hat{T} \). The set of feasible solutions is obtained by using both the model assumptions and the identification conditions. In particular, the set of feasible transportation costs is given by:

\[
\hat{T} = \left\{ \hat{\tau} \in \mathbb{R}_{++} \mid \lambda^t(\hat{\tau}) \in [0, 1] \quad \forall t \quad \land \quad \nu^t(\hat{\tau}) \geq 0 \quad \forall t \quad \land \quad \hat{\alpha}^{\text{nls}}(\hat{\tau}) \geq 1 \quad \land \quad \hat{\beta}^{\text{nls}}(\hat{\tau}) \geq 1 \right\}
\]

where the first condition requires that the domain of the function must lie between 0 and 1 – which is the domain of the beta distribution; the second condition guarantees the non-negativity of the range of the distribution of consumers; and the third and fourth conditions are imposed to satisfy Assumption 1 (log-concave distribution function).

Therefore, the estimation of the transportation cost (\( \hat{\tau} \)) is reduced to a local optimization problem, which can be written as follows:

\[
\hat{\tau} \equiv \arg \min_{\tau \in \hat{T}} \sum_{t=0}^{T} \left[ \nu^t(\tau) - f(x^t, \hat{\alpha}^{\text{nls}}(\hat{\tau}), \hat{\beta}^{\text{nls}}(\hat{\tau})) \right]^2
\]

(5.17)

If there exist at least one set of identified \( \lambda \)'s and \( \nu \)'s that lies in the feasible
solutions set and the estimated shape parameters are greater than or equal to 1 for every candidate \( \tau \), \( \hat{T} \) is a closed and bounded set.\(^{21}\) Therefore, the extreme value theorem guarantees the existence of a solution.

To give an idea about the performance of this estimator, Table 5.1 includes the estimated transportation costs—as well as the estimation of the shape parameters—for a simulation study based on 2,000 observations using the following numbers (we use the same parameters that we employ in Appendix F). First, firms’ marginal costs at each period are drawn from a uniform distribution between 0.7 and 1.2. That is \( c(\theta^t) \sim U[0.7, 1.2] \) for all \( t \). Second, firms’ locations are such that \( x_1 = \frac{1}{3} \) and \( x_2 = \frac{2}{3} \), with the interval normalized to \([0, 1]\). Third, the (known) consumers’ transportation cost is \( \tau = 1.5 \). As we can see, the estimation of the transportation cost is very close to the real transportation cost in the three beta distributions considered.\(^{22}\)

Table 5.1: Transportation cost and shape parameters for different Beta distributions \((n = 2000)\)

<table>
<thead>
<tr>
<th>( \tau = 1.5 )</th>
<th>( \alpha = 2.5, \beta = 5 )</th>
<th>( \alpha = 2, \beta = 2 )</th>
<th>( \alpha = 4.5, \beta = 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\tau} )</td>
<td>1.50012</td>
<td>1.50007</td>
<td>1.50012</td>
</tr>
<tr>
<td>( \hat{\alpha}^{nls} )</td>
<td>2.50049</td>
<td>2.00017</td>
<td>4.50067</td>
</tr>
<tr>
<td>( \hat{\beta}^{nls} )</td>
<td>5.00082</td>
<td>2.00017</td>
<td>3.00050</td>
</tr>
</tbody>
</table>

5.4 The Spanish retail gasoline market data

In the present section we apply our methodology to a dataset that consists of prices and costs of a pair of gas stations located along a straight highway in Spain. We organize this section as follows. First, we briefly describe the main characteristics of

\(^{21}\)In particular, using the conditions above, the set is bounded below by \( \max\left\{ \min\left\{ 0, \frac{\nu_{1}^* - \nu_{2}^*}{(x_2 - x_1)^2} \right\}, \min\left\{ 0, \frac{\nu_{1}^* - \nu_{2}^*}{(x_2 - x_1)(2 - x_2 - x_1)} \right\} \right\} \) and the set can be bounded above (if necessary) considering that the sum of the (unique) \( \nu \)'s must be less than or equal to 1—which is a requirement of the distribution itself—, i.e. \( \tau \leq \sum \frac{\nu_{1}^* - c(\theta^t) + \nu_{2}^* - c(\theta^t)}{2(x_2 - x_1)} \) for every unique pair of equilibrium prices and costs.

\(^{22}\)We obtain very similar results using the least absolute deviations estimator.
the retail gasoline market in Spain (subsection 5.4.1). Then, we present the data and some descriptive statistics (subsection 5.4.2). Finally, in subsection 5.4.3, we estimate the distribution of consumers using our dataset and we conduct some counterfactual welfare experiments.

5.4.1 Main features of the market

As we mentioned above, there are several reasons that make the Spanish retail gasoline market an appropriate one to use our methodology. To begin with, the geographic conditions of the Iberian Peninsula are appropriate to build straight roads. Second, the retail Spanish gasoline market is a heavily concentrated one. In fact, as Bello and Contín-Pilart (2012) indicate, just five major operators controlled about 75-80% of the market in 2008 –see also Contín-Pilart et al. (2009).

This lack of competition in highways, which was previously pointed out by Contín-Pilart and Pintado (2010), is translated into oligopolistic behavior in highways and greater markups. Some evidence of this fact is provided in Table 5.2, which shows the coefficient obtained from regressing the retail price of the fuels on several variables, including a dummy for those gas stations located along the selected toll highway. For both types of fuel considered –regular unleaded gasoline and regular diesel fuel– the toll highway dummy coefficient is positive and statistically significant.23

Third, the typical contracts between gas stations (retailers) and wholesale fuel sellers are very standard and widely used. In particular, wholesalers sell gasoline and diesel to retailers at a price that is determined by Platts24 daily indices plus a per-
Table 5.2: The “highway premium”: evidence of higher prices in highways

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>log(price unleaded gas)</th>
<th></th>
<th>log(price diesel)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RE</td>
<td>H-T</td>
<td>RE</td>
<td>H-T</td>
</tr>
<tr>
<td>Highway</td>
<td>0.037***</td>
<td>0.037***</td>
<td>0.050***</td>
<td>0.050***</td>
</tr>
<tr>
<td></td>
<td>(0.0038)</td>
<td>(0.0058)</td>
<td>(0.0048)</td>
<td>(0.0080)</td>
</tr>
<tr>
<td>log(Brent)</td>
<td>0.270***</td>
<td>0.270***</td>
<td>0.299***</td>
<td>0.299***</td>
</tr>
<tr>
<td></td>
<td>(0.0011)</td>
<td>(0.0007)</td>
<td>(0.0011)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>24/7</td>
<td>-0.012**</td>
<td>-0.011***</td>
<td>-0.014**</td>
<td>-0.013***</td>
</tr>
<tr>
<td></td>
<td>(0.0056)</td>
<td>(0.0008)</td>
<td>(0.0067)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>Major brand</td>
<td>0.006**</td>
<td>0.005***</td>
<td>0.004</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>(0.0031)</td>
<td>(0.0005)</td>
<td>(0.0042)</td>
<td>(0.0005)</td>
</tr>
</tbody>
</table>

| Month dummy         | Yes                     | Yes | Yes               | Yes |
|                     |                         |     |                   |     |
| Day of the week dummy | Yes                   | Yes | Yes               | Yes |
| Observations        | 369,455                 | 369,455 | 382,039         | 382,039 |
| # of gas stations   | 1,353                   | 1,353 | 1,401            | 1,401 |

RE indicates Random Effects and H-T indicates Hausman-Taylor
Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

The “highway premium”: evidence of higher prices in highways

liter markup. There are two major indices that set wholesale gasoline prices. First, and most importantly for the Spanish case, the Genoa market in Italy (also known as Mediterranean or MED market). In particular, the reference index in this market for unleaded gasoline is the “Premium Unleaded 10 ppm MED CIF Cargoes Mid” index, while for diesel fuel is the “ULSD 10 ppm MED CIF Cargoes Mid”. Second, the Rotterdam market in the Netherlands (also known as North-Western Europe or NWE market) is also a relevant one. In particular, for unleaded gasoline the key index is the “Gasoline 10 ppm NWE CIF ARA”, while for diesel fuel the key index is the “ULSD 10 ppm NWE CIF Cargoes Mid”.

Gas stations’ owners adjust the price of unleaded gasoline and diesel fuel (the most commonly sold fuels in Spain) every day according to the cost index in the Genoa and Rotterdam markets posted in the Platts. Moreover, every day each gas station must send to the Ministry of Industry the price at which they are selling their fuels and agriculture information, and a premier source of benchmark price assessments for those commodity markets”. For more information, see http://www.platts.com/.
the following day.\textsuperscript{25} The Ministry of Industry makes all gas stations’ prices publicly available every day\textsuperscript{26} and erases the previous day prices.

5.4.2 Data and descriptive statistics

Our dataset includes daily prices from October 2, 2014 to July 20, 2015\textsuperscript{27}. Data on prices is available for the most commonly used fuels in Spain for transportation purposes, namely unleaded gasoline (95 RON) and diesel A (regular). These prices were downloaded every day from the Ministry of Industry website.\textsuperscript{28}

Posted prices obtained from the Ministry of Industry are tax-inclusive. Therefore, we need to subtract the taxes included in these prices.\textsuperscript{29} In particular, we need to subtract the VAT –which is determined at the national level– and the excise taxes –which are determined both at the national and at the regional level. In particular, the state tranche of the excise duty is 0.42469€/l for unleaded gasoline and 0.331€/l for diesel fuel. The regional tranche varies from one region to another. The VAT base is the retail (tax-exclusive) price plus the excise duties. The VAT rate is 21%.

Data on costs for the same period (from October 2, 2014 to July 20, 2015) was obtained from \textit{Platts}. Using data from both the Genoa market and the Rotterdam market, we build a representative index of the variable costs that face each gas station. Thus, according to the information provided by the Spanish Association of Operators of Oil Products (AOP)\textsuperscript{30} and the National Competition Commission (CNC), the

\textsuperscript{25}Further information about this procedure to send prices is in the Spanish Act ITC/2308/2007.
\textsuperscript{26}The data is available at \url{http://geoportalgasolineras.es/descargas.do?tipoBusqueda=0}.
\textsuperscript{27}Data is not available for October 12, 2014, which is the Spanish National day.
\textsuperscript{28}\url{http://geoportalgasolineras.es/descargas.do?tipoBusqueda=0}.
\textsuperscript{29}Marion and Muehlegger (2011) show that gasoline taxes are indeed fully passed onto consumers and are incorporated fully into the tax-inclusive price, under typical supply and demand conditions.
\textsuperscript{30}Both Shell and Cepsa are partners of this organization.
representative variable cost of unleaded gasoline for a retailer is calculated as follows\textsuperscript{31}:

\[
\text{VC Unleaded Gasoline} = 70\% \times \text{“Premium Unleaded 10 ppm MED CIF Cargoes Mid” Genoa index} + \\
+ 30\% \times \text{“Gasoline 10 ppm NWE CIF ARA” Rotterdam index}
\]

Likewise, the variable cost of the diesel fuel for a retailer is equal to

\[
\text{VC Diesel Fuel} = 70\% \times \text{“ULSD 10 ppm MED CIF Cargoes Mid” Genoa index} + \\
+ 30\% \times \text{“ULSD 10 ppm NWE CIF Cargoes Mid” Rotterdam index}
\]

The weighted average costs obtained from \textit{Platts} are in \$/metric ton. To convert from metric tons to liters, we use the reference density of unleaded gasoline (around 0.745 kg/l), and regular diesel fuel for vehicles (around 0.850 kg/l). We get a rate of 1315.78 l/metric ton and 1183.43 l/metric ton respectively. To convert from \$ to €, we use the daily exchange rate from October 2, 2014 to July 20, 2015. This data can be obtained from the Federal Reserve Bank of St. Louis and is publicly available.\textsuperscript{32,33}

To empirically test our procedure, we consider data from a pair of gas stations located in a 35.1 miles (56.5 km) straight section in the AP-7 Toll Highway in Catalonia. This straight section of the highway is located between Amposta and Cambrils. Figure 5.2 shows this section of the road and indicates the position of the two gas stations. Point A indicates the former town, while point B indicates the latter one.

Gas station #1 is located 3.1 miles away from Amposta, close to the village of

\textsuperscript{31}This formula is used by the Spanish Association of Operators of Oil Products to calculate the variable cost that gas stations face in Spain. The formula is included in their monthly reports. E.g. February 2015 report (http://www.aop.es/informes/informes_sector/composicion-delprecio-0215.pdf). Likewise, the Spanish National Competition Commission uses the same formula to calculate gas stations' markups. See, for instance, the 2012 special report on the gasoline market (http://www.cmnc.es/Portales/0/Ficheros/Promocion/Informes_y_Estudios_Sectoriales/2012/Informe%20Distribucion%20de%20Carburantespdf.pdf, p. 58). A similar weighted average of the relative importance of Genoa and Rotterdam prices (66.1% the former, 33.9% the latter) was employed by Rodrigues (2009) to study asymmetries in the adjustment of pump prices for the Spanish case.

\textsuperscript{32}https://research.stlouisfed.org/fred2/series/DEXUSEU.

\textsuperscript{33}We also include as variable cost the wholesalers' markup, which is approximately 0.015€/l according to the information provided by the AOP.
L’Aldea. Between Amposta and gas station #1 the highway gives access to roads N-235 and N-340, which connect several towns along the highway, such as L’Aldea. Gas station #2 is located 24.11 miles away from Amposta, close to the city of L’Hospitalet de l’Infant. Between gas station #1 and gas station #2 the highway has a few exits to several towns, such as Calafat, Les Tres Cales, and another exit to the National Park of La Rojala-Platja del Torn. Finally, from gas station #2 to Cambrils, the highway gives access to roads C-44 and T-312. These roads connect popular beaches and holiday resorts in Salou and Cambrils, and a well-known theme park called Port Aventura.

We normalize the interval defined by the section of the highway in between Amposta and Cambrils to $[0,1]$. Following such normalization, gas station #1 is located at point $x_1 = 0.09$ and gas station #2 is located at point $x_2 = 0.68$. Both gas stations are operated by the same major brand (Cepsa), are opened 24/7 and have a
restaurant and a convenience store. The regional tranche in Catalonia for both fuel types is 0.048€/l.

Summary statistics for the relevant variables from October 2, 2014 to July, 2015 are included in Table 5.3. All the variables are expressed in euros per liter. As we can see, all the variables present substantial variation. In fact, during the last quarter of 2014 oil prices worldwide experimented a huge decrease. This scenario is convenient for our analysis since this pattern was also observed in the Genoa and Rotterdam prices, creating variation in gas stations’ costs and posted prices.

Table 5.3: Summary statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prices, tax-inclusive (€/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unleaded gasoline, gas station #1</td>
<td>1.3242</td>
<td>0.0745</td>
<td>1.1690</td>
<td>1.4690</td>
</tr>
<tr>
<td>Unleaded gasoline, gas station #2</td>
<td>1.3283</td>
<td>0.0764</td>
<td>1.1690</td>
<td>1.4690</td>
</tr>
<tr>
<td>Diesel fuel, gas station #1</td>
<td>1.2374</td>
<td>0.0611</td>
<td>1.1090</td>
<td>1.3790</td>
</tr>
<tr>
<td>Diesel fuel, gas station #2</td>
<td>1.2388</td>
<td>0.0606</td>
<td>1.1090</td>
<td>1.3790</td>
</tr>
<tr>
<td>Prices, tax-exclusive (€/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unleaded gasoline, gas station #1</td>
<td>0.6217</td>
<td>0.0616</td>
<td>0.4934</td>
<td>0.7414</td>
</tr>
<tr>
<td>Unleaded gasoline, gas station #2</td>
<td>0.6251</td>
<td>0.0632</td>
<td>0.4934</td>
<td>0.7414</td>
</tr>
<tr>
<td>Diesel fuel, gas station #1</td>
<td>0.6437</td>
<td>0.0505</td>
<td>0.5375</td>
<td>0.7607</td>
</tr>
<tr>
<td>Diesel fuel, gas station #2</td>
<td>0.6448</td>
<td>0.0501</td>
<td>0.5375</td>
<td>0.7607</td>
</tr>
<tr>
<td>Costs data (€/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unleaded gasoline</td>
<td>0.4268</td>
<td>0.0596</td>
<td>0.2952</td>
<td>0.5354</td>
</tr>
<tr>
<td>Diesel fuel</td>
<td>0.4389</td>
<td>0.0486</td>
<td>0.3324</td>
<td>0.5462</td>
</tr>
</tbody>
</table>

Observations per variable 291

5.4.3 Main results and counterfactuals

In this subsection we use the estimation procedure described in the previous section to estimate the distribution of consumers using the AP-7 Toll Highway data. Base on this estimation, we calculate the socially optimal location of an entrant gas station (i.e. the location that maximizes consumers’ welfare).

The main results of the estimation analysis are included in Figures 5.3 and ??.

Subfigure 5.3a contains the estimated distribution of consumers for the section of the
AP-7 Toll Highway. As we can see, the distribution of consumers is right-skewed for both unleaded gasoline consumers and diesel consumers. This result is consistent with the fact that average prices of unleaded gasoline and diesel fuel in gas station #2—the one that is to the right, i.e. the closest to Cambrils (town B)—are higher than average fuel prices in gas station #1. Moreover, we also know that gas station #2 is close to very popular tourist attractions and landmark, such as beach resorts, a theme park and a national park. Therefore, consistently with the estimation, we should expect more traffic in the section of the highway that is closer to Cambrils (town B) than in the section of the highway that is closer to Amposta (town A).

Figure 5.3: Estimated distribution of consumers

![Graph showing distribution of consumers for unleaded gasoline and diesel fuel](image)

(a) Unleaded gasoline  (b) Diesel fuel

In fact, as we can see in Figure 5.4, the estimated distribution of consumers and, more precisely, the peak of the distribution obtained, coincides with the typically congested section of this highway. That is, our estimated distribution is also consistent with the congestion of the selected section of the AP-7 highway.

Subfigure 5.5 includes the optimal location of an entrant gas station. Consistently with the estimated distribution of consumers, the optimal location point of an entrant is at point $x = 1$, i.e. close to Cambrils. Thus, our model suggests that a new gas
Figure 5.4: Map of the section of the AP-7 Toll Highway

station serving the popular tourist areas is the optimal location in terms of welfare.
If a new gas station is to be located in between gas station #1 and gas station #2,\textsuperscript{34} our estimation using both types of fuels suggests that the optimal location is at point $x = 0.38$. This point is located 13.4 miles away from Amposta (town A) and just close to the “Estany podrit” beach and the Ametlla camping. As expected, this point is slightly closer to gas station #2 than to gas station #1.\textsuperscript{35}

Figure 5.5: Incumbents’ locations and entrant’s socially optimal location

Finally, using the estimated distribution of consumers and transportation costs, we

\textsuperscript{34}We call a point between gas station #1 and gas station #2 a mid-point location.

\textsuperscript{35}Since the results for both fuels are very similar, we condense the results of both fuels in just one diagram.
study the welfare gains from locating an entrant gas station in the estimated socially optimal locations in comparison to locations based purely on geographical distances, i.e. without taking into account the underlying distribution of consumers. The main results of this welfare counterfactual exercise are included in Table 5.4.

Table 5.4: Percent of welfare gains from optimal location of gas stations vs. geographical mid-point locations

<table>
<thead>
<tr>
<th>Location Description</th>
<th>Socially optimal location (x=1)</th>
<th>Socially optimal mid-point location (x=0.38)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid-point between town A &amp; gas station #1 (x=0.045)</td>
<td>1.48%</td>
<td>0.20%</td>
</tr>
<tr>
<td>Mid-point between both gas stations (x=0.355)</td>
<td>1.14%</td>
<td>0.03%</td>
</tr>
<tr>
<td>Mid-point between gas station #2 &amp; town B (x=0.810)</td>
<td>0.05%</td>
<td>-0.49%</td>
</tr>
</tbody>
</table>

As indicated, there are welfare gains from locating an entrant gas station in the socially optimal location in comparison to the three proposed geographical mid-point locations. As expected, these welfare gains are lower as the alternative locations get closer to the optimal one. For instance, in the AP-7 case, locating the entrant firm at the optimal location versus locating it at the geographical mid-point between town A and gas station #1 increases welfare by 1.48%. However, locating the entrant firm at the optimal point in comparison to the geographical mid-point between gas station #2 and town B increases welfare just by 0.05%. The reason is that the latter point is closer to the optimal point than the former one.

Column 2 in Table 5.4 includes welfare losses from locating an entrant gas station

---

36 We have considered as “locations based on geographical distances” for an entrant firm the following ones. First, the geographical mid-point between town A and gas station #1. Second, the geographical mid-point between gas station #1 and gas station #2. Third, the geographical mid-point between gas station #2 and town B.

37 The results included in this table are preliminary.
in the optimal point in between both gas stations in both highways (the socially optimal mid-point location). The results suggests that it is welfare enhancing for consumers to locate an entrant at $x = 0.81$ –i.e. between gas station #2 and town B– than to locate it at point $x = 0.38$ –the socially optimal mid-point location. Again, the reason is that $x = 0.81$ is closer than $x = 0.38$ to the absolute socially optimal location, which is $x = 1$.

### 5.5 Conclusions

This chapter proposes a novel procedure to recover the (potentially non-uniform) distribution of consumers in an oligopolistic market serving horizontally differentiated products –i.e. a Hotelling linear city model– with exogenously given firms’ locations. The procedure relies on sufficient variation in equilibrium prices and variable costs from one period to another. By imposing some parametric restrictions, our methodology allows us also to recover a precise approximation of the usually called transportation cost.

This methodology is applied to data on prices and costs of a pair of gas stations located along a straight highways in Spain. By doing so, we recover the underlying distribution of consumers along straight sections of main highways between two towns. Using the estimated distribution and transportation cost, we are able to precisely indicate where should be located a potential entrant to maximize consumers’ welfare. Opening a gas station in a Spanish highway requires an authorization issued by the Ministry of Public Works and Transport about the convenience of the potential location. Hence, our methodology aims to add some light to this opaque bureaucratic process.

Besides the highway application, this methodology also has numerous potential applications. For instance, we can use it to determine the optimal position of oil and
gas storage facilities along straight pipelines.

Coming back to the gas stations case, possible extensions should take into account the potential existence of collusive behavior. This fact is pointed out for the case of gas stations by Borenstein and Shepard (1996). In addition to that, further work should consider price inertia or rigidities in prices and costs. Although these elements do not apply to the retail gasoline market, they can play a prominent role in other markets in which we observe spatial competition or horizontal differentiation.
Chapter 6

Thesis conclusions

In this thesis, I discuss different issues related to energy policy, particularly related to energy access, supply reliability and efficiency (welfare). I study these issues in the context of two particular energy markets: the electricity markets and gasoline retail markets.

Even though these look like distinct markets, the connection between them has been rapidly increasing. The current concern about climate change and environmental anomalies, which were translated into several important mandates, has resulted in an increasing switching from gasoline-powered vehicles to electricity-powered vehicles. In a context with a rapid increase of the electric vehicle fleet, electricity access, capacity generation adequacy, and fuel policy must be coordinated to address several challenges such as the rapid increase in electricity demand and the (potential) decrease in gasoline and diesel usage to power vehicles.

While there is a connection between these two markets, this thesis does not address that issue. Rather it focuses on important issues relevant only to one market or the other. Nevertheless, the issues we do address in this thesis will play a role in considerations about what to do as these markets become more integrated.

Chapter 1 discussed the role of capacity markets as a mechanism to solve the
so-called “missing money” (or resource adequacy) problem. Using a stylized competitive model, we provide the conditions under which a capacity market improves market efficiency. Chapter 2 takes the theoretical exercise to the particular case of the Texas ERCOT market. We use data on costs and hourly load to evaluate the equilibrium market conditions previously obtained. We find that a capacity market solves the reliability problem and reduces the volatility of electricity price at the cost of increasing average consumers’ electricity bills.

Chapter 3 studies the impact of changes in taxes and prices on consumers’ demand for fuels used for transportation purposes. Using some standard panel data techniques, we find that the elasticity of demand in response to changes in taxes is much greater than the elasticity of demand in response to changes in (tax-exclusive) gasoline prices.

Chapter 4 studies a gas-station duopoly game of spatial competition in prices. Using the equilibrium conditions, we are able to extract some information about consumers’ demand. In particular, differences in prices reveal the pattern of traffic distribution in different sections of major roads. We show how this provides a test for a standard model used in transportation economics. After validating the model, we show how it can be used to find optimal location of entrant firms.
Bibliography


Appendices
Appendix A

**Proof of Proposition 1**: given both electricity generators capacities, \( k_i \in [0, 1], \ i \in \{b, p\} \), we face three possible scenarios:

1. \( k_b \geq \theta \). In this case, we claim that the bidding strategy for the base load generator is such that \( p_b = c_b \) and the bidding strategy for the peak load generator is such that \( p_p \in (p_b, \infty) \). To prove this fact, let us consider possible deviations for generator \( i \) given generator \( j \)'s strategy, \( i, j \in \{b, p\}, i \neq j \). Let us begin with the potential base load generator deviations: first, given that \( p_p \in (p_b, \infty) \), let us assume that \( p_b > c_b \). Then a base load generator can choose an arbitrary small \( \epsilon > 0 \) such that \( p_b > p_b - \epsilon > c_b \), and this generator makes a positive profit. Thus, the proposed deviation cannot be an equilibrium (the base load generators will deviate undercutting \( p_b \) until \( p_b = c_b \)). Second, given that \( p_p \in (p_b, \infty) \), let us assume that \( p_b < c_b \). Then, the base load generator serves the whole demand. But notice that this generator makes negative profit (since in this case it sells each unit of electricity at a price lower than the variable cost of producing it). I.e. \( \pi_b < 0 \), which cannot be true in equilibrium. Next, let us deal with potential peak load generator deviations: given that \( p_b = c_b \), let us assume that \( p_p \leq p_b \). Then, the peak load generator serves at least half of the demand. But notice that \( p_p \leq p_b = c_b < c_p \iff p_p < c_p \), which implies that this generator makes a negative profit (since in this case it sells each unit of electricity at a price lower than the variable cost of producing it). I.e. \( \pi_p < 0 \), which cannot be true in equilibrium. Thus, if \( k_b \geq \theta \), in equilibrium the base load generator’s bid always clears the auction. Hence, the equilibrium price in the wholesale spot market is \( p^s = p_b = c_b \) and all the demand is served by the base load generator. Moreover, since this generator sells at price equal marginal cost, it makes no profit \( \pi_b = 0 \). Of course, since the peak generator does not produce and sell, \( \pi_p = 0 \).

2. \( k_b < \theta \leq K \). In this case, we claim that the bidding strategy for the base load generator is such that \( p_b \in (0, p_p) \) and the bidding strategy for the peak load generator is such that \( p_p = c_p \). To prove this fact, let us consider again possible deviations for both electricity generators. To begin with, let us consider base load generator deviations: given that \( p_p = c_p \), let us assume that \( p_b = p_p \);
then, instead of serving \( k_b \) units of electricity, the base load generator serves 
\[
\frac{1}{2} \min \{ \theta, k_b \} = \frac{1}{2} k_b.
\]
Thus, this deviation is not profitable for this generator. Next, assume that \( p_b > p_p \). If \( \theta \le k_p \), the peak load generator serves the whole demand, and the base load generator does not serve, so the base load generator is worse off. If \( \theta > k_p \), then the base load generator serves a positive amount. But notice that a base load generator can choose an arbitrarily small \( \epsilon > 0 \) such that \( p_b > p_b - \epsilon > p_p \). Thus, by offering \( p_b - \epsilon \), this generator captures the whole residual demand, and hence the base load generators will deviate undercutting \( p_b \) until \( p_b = p_p \). Next, let us deal with potential peak load generator deviations: first, given that \( \theta \le k_p \), in equilibrium the peak load generator’s bid always clears the auction. Hence, the equilibrium price in the wholesale spot market is \( p^s = p_p = c_p \) and the base load generator serves \( k_b \), and the peak load generator serves the residual demand \( \theta - k_b \). Moreover, since the peak load generator sells at price equal marginal cost, it makes no profit \( \pi_p = 0 \). On the other hand, the base load generator sells at price greater than its marginal cost. Thus, its profit is given by \( \pi_b = c_p k_b - c_b k_b \).

3. \( K < \theta \) (part of the demand suffers a blackout). In this case, we claim that the bidding strategy for both generators is such that \( p_b = p_p = p^H \). To prove this fact, let us consider again possible deviations. First, fix \( p_j = p^H \) and assume that \( p_i < p^H \), for \( i, j \in \{ b, p \}, i \neq j \). Then a consumer suffering a blackout (i.e. with zero surplus) can choose an arbitrary small \( \epsilon > 0 \) such that \( p_i < p_i + \epsilon < p^H \), and such consumer will obtain positive surplus by paying \( p_i + \epsilon \). Thus, the proposed deviation cannot be an equilibrium, since consumers suffering the blackout will deviate by reversely-undercutting \( p_i \) until \( p_i = p^H \). Second, fix \( p_j = p^H \) and assume that \( p_i > p^H \), for \( i, j \in \{ b, p \}, i \neq j \). Then, the generator \( i \) does not serve the demand, since the consumers are not willing to pay more than \( p^H \). Thus, this deviation is not profitable for generator \( i \), so it cannot be an equilibrium. Thus, if \( K < \theta \), the equilibrium price in the wholesale spot market is \( p^s = p^H \) and both generators produce at maximum capacity \( (q_i = k_i, i \in \{ b, p \}) \) and they make profits such that \( \pi_i = p^H k_i - c_i k_i, i \in \{ b, p \} \).
Proof of Proposition 2: let us begin with the type-1 firms; denote $\pi_b$ type-1 firms’ aggregate profit and $\mathbb{E}$ the expectation operator. Then, the expected profit for them is:

$$\mathbb{E}\pi_b = \int_0^{k_b} 0 q_b dF(\theta) + \int_{k_b}^K (c_p - c_b) q_b dF(\theta) + \int_K^1 (p^H - c_b) q_b dF(\theta) - c_{k_b} k_b$$

s.t. $0 \leq q_b \leq k_b$

We can simplify the previous expression as follows:

$$\mathbb{E}\pi_b = \int_{k_b}^K (c_p - c_b) k_b dF(\theta) + \int_K^1 a_{p_b} q_b dF(\theta) - c_{k_b} k_b$$

where $a_{b} \equiv p^H - c_b$. Notice that when $k_b \leq \theta$ the type-1 firms produce at maximum capacity. Therefore, in equilibrium, $q_b = k_b$ whenever $k_b \leq \theta$. Thus:

$$\mathbb{E}\pi_b = \int_{k_b}^K (c_p - c_b) k_b dF(\theta) + \int_K^1 a_{b} k_b dF(\theta) - c_{k_b} k_b$$

Solving for the integrals, we arrive at:

$$\mathbb{E}\pi_b = (c_p - c_b) k_b [F(K) - F(k_b)] + a_{b} k_b [1 - F(K)] - c_{k_b} k_b$$

In equilibrium, since we assume free entry and perfect competition, the expected profit must be equal to zero. Thus:

$$(c_p - c_b) k_b [F(K) - F(k_b)] + a_{b} k_b [1 - F(K)] - c_{k_b} k_b = 0$$

Rearranging, we arrive at:

$$(c_p - c_b) F(k_b) = a_b - a_p F(K) - c_{k_b} \quad \text{(A.1)}$$

Now let us turn to the type-2 firms; denote $\pi_p$ type-2 firms’ aggregate profit. Thus:

$$\mathbb{E}\pi_p = \int_0^{k_p} 0 q_p dF(\theta) + \int_{k_p}^K 0 q_p dF(\theta) + \int_K^1 (p^H - c_p) q_p dF(\theta) - c_{k_p} k_p$$

s.t. $0 \leq q_p \leq k_p$

We can simplify the previous expression as follows:

$$\mathbb{E}\pi_p = \int_K^1 a_{p} q_p dF(\theta) - c_{k_p} k_p$$

s.t. $0 \leq q_p \leq k_p$
where \(a_p \equiv p^H - c_p\). Notice that when \(K \leq \theta\) the type-2 firms produce at maximum capacity. Therefore, in equilibrium, \(q_p = k_p\) whenever \(K \leq \theta\). Thus:

\[
\mathbb{E}\pi_p = \int_K^1 a_p k_p dF(\theta) - c_k p k_p
\]

Solving for the integrals, we arrive at:

\[
\mathbb{E}\pi_p = a_p k_p [1 - F(K)] - c_k p k_p
\]

In equilibrium, since we assume free entry and perfect competition, the expected profit must be equal to zero. Thus:

\[
a_p k_p [1 - F(K)] - c_k p k_p = 0
\]

Rearranging, we arrive at:

\[
F^{-1}\left(1 - \frac{c_k p}{a_p}\right) = k_p \tag{A.2}
\]

Plugging equation A.2 into equation A.1:

\[
(c_p - c_b) F(k_b) = a_b - a_p F\left[F^{-1}\left(1 - \frac{c_k p}{a_p}\right)\right] - c_k b
\]

Solving for \(k_b\) we arrive at:

\[
k_b = F^{-1}\left(1 - \frac{c_k b - c_k p}{c_p - c_b}\right) \tag{A.3}
\]

Plugging equation A.3 into equation A.2 and solving for \(k_p\) we arrive at:

\[
k_p = F^{-1}\left(1 - \frac{c_k p}{a_p}\right) - k_b \tag{A.4}
\]

Therefore, in equilibrium equations A.3 and A.4 must hold. Let us first deal with potential corner solutions in a case-by-case basis:

- Case #1: \(k_b \leq 0\) and \(k_p \leq 0\). Thus, by equation A.3: \(k_b = F^{-1}\left(1 - \frac{c_k b - c_k p}{c_p - c_b}\right) \leq 0 \iff 1 - \frac{c_k b - c_k p}{c_p - c_b} \leq 0 \iff 1 - \frac{c_k b - c_k p}{c_p - c_b} \leq 0 \iff c_p - c_b \leq c_k b - c_k p\) and by equation A.4 \(k_p = F^{-1}\left(1 - \frac{c_k p}{a_p}\right) \leq 0 \iff 1 - \frac{c_k p}{a_p} \leq 0 \iff c_k p \geq a_p\).

- Case #2: \(k_b \leq 0\) and \(k_p \geq 1\). Thus by equation A.4: \(k_p = F^{-1}\left(1 - \frac{c_k p}{a_p}\right) \geq 1 \iff 1 - \frac{c_k p}{a_p} \geq 1 \iff \frac{c_k p}{a_p} \geq 0\), which is a contradiction, since \(c_k p > 0\) and \(a_p > 0\).

- Case #3: \(k_b \geq 1\) and \(k_p \leq 0\). Thus by equation A.3: \(k_b = F^{-1}\left(1 - \frac{c_k b - c_k p}{c_p - c_b}\right) \geq 1 \iff 1 - \frac{c_k b - c_k p}{c_p - c_b} \geq 0 \iff c_p - c_b \geq c_k b - c_k p\).
1 \iff 1 - \frac{c_{kb} - c_{kp}}{c_p - c_b} \geq 1 \iff -\frac{c_{kb} - c_{kp}}{c_p - c_b} \geq 0 \text{ which is a contradiction, since } c_{kb} > c_{kp} \text{ and } c_p > c_b.

• Case #4: $k_b \geq 1$ and $k_p \geq 1$. Thus by equation A.3: $k_b = F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right) \geq 1 \iff 1 - \frac{c_{kb} - c_{kp}}{c_p - c_b} \geq 1 \iff -\frac{c_{kb} - c_{kp}}{c_p - c_b} \geq 0 \text{ which is a contradiction, since } c_{kb} > c_{kp} \text{ and } c_p > c_b.$

• Case #5: $k_b \leq 0$ and $k_p \in (0,1)$ (interior). Thus by equation A.4: $k_p = F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right)$ and notice that $k_p \in (0,1) \iff a_p > c_{kp}$. For type-1 firm, following equation A.3, we require: $k_b = F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right) \leq 0 \iff 1 - \frac{c_{kb} - c_{kp}}{c_p - c_b} \leq 0 \iff 1 \leq \frac{c_{kb} - c_{kp}}{c_p - c_b} \iff c_p - c_b \leq c_{kb} - c_{kp}.$

• Case #6: $k_b \in (0,1)$ (interior) and $k_p \leq 0$. Thus by equation A.3: $k_b = F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right)$ and notice that $k_b \in (0,1) \iff c_{kb} - c_{kp} < c_p - c_b$. For the type-2 firms, following equation A.4, we require $k_p = F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right) - F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right) \leq 0 \iff F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right) \leq F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right) \iff 1 - \frac{c_{kp}}{a_p} \leq 1 - \frac{c_{kb} - c_{kp}}{c_p - c_b} \iff \frac{c_{kp}}{a_p} \geq \frac{c_{kb} - c_{kp}}{c_p - c_b} \iff c_{kp}(a_p - c_{kp}) \geq a_p(c_p - c_{kb}) \iff c_{kp}a_b - c_{kp}a_p \geq a_p(c_{kb} - c_{kp}) \iff c_{kp}a_b \geq a_p(c_{kb} - c_{kp}).$

• Case #7: $k_b \geq 1$ and $k_p \in (0,1)$ (interior). Thus by equation A.3: $k_b = F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right) \geq 1 \iff 1 - \frac{c_{kb} - c_{kp}}{c_p - c_b} \geq 1 \iff -\frac{c_{kb} - c_{kp}}{c_p - c_b} \geq 0 \text{ which is a contradiction, since } c_{kb} > c_{kp} \text{ and } c_p > c_b.$

• Case #8: $k_b \in (0,1)$ (interior) and $k_p \geq 1$. Thus by equation A.3: $k_b = F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right)$ and notice that $k_b \in (0,1) \iff c_{kb} - c_{kp} < c_p - c_b$. For type-2 firms, following equation A.4, we require: $k_p = F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right) - k_b \geq 1 \iff F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right) \geq 1 + k_b$. Since $k_b \in (0,1)$, then: $F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right) \geq 1 + k_b \Rightarrow F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right) > 1 \iff 1 - \frac{c_{kp}}{a_p} > 1 \iff -\frac{c_{kp}}{a_p} \geq 0 \text{ which is a contradiction, since } c_{kb} > 0 \text{ and } a_p > 0.$

Finally, the interior solution ($k_i \in (0,1)$ for $i \in \{b, p\}$) is given by equations A.3 and A.4:

\[
k_b = F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right)
\]

and
\[
k_p = F^{-1} \left(1 - \frac{c_{kp}}{a_p}\right) - F^{-1} \left(1 - \frac{c_{kb} - c_{kp}}{c_p - c_b}\right)
\]

for the remaining cases. \qed
Proof of Corollary 1: 

\[ k_b^* = F^{-1}\left(1 - \frac{ck_b - c_k}{c_p - c_b}\right) \quad \text{and} \quad k_p^* = F^{-1}\left(1 - \frac{ck_p}{a_p}\right) - F^{-1}\left(1 - \frac{ck_b - c_k}{c_p - c_b}\right) \]

which implies that \( K^* = F^{-1}\left(1 - \frac{ck_b - c_k}{c_p - c_b}\right) + F^{-1}\left(1 - \frac{ck_p}{a_p}\right) - F^{-1}\left(1 - \frac{ck_b - c_k}{c_p - c_b}\right) \). I.e. \( K = F^{-1}\left(1 - \frac{ck_p}{a_p}\right) \). And since \( a_p > c_k \) (see proof of Proposition 2), \( K^* \in (0, 1) \).  

Proof of Proposition 3: the proof is similar to the proof of Proposition 2, except that there is a compensation equal to \( m \) per unit of capacity built and that \( K^T \) denotes total invested capacity.  

Proof of Corollary 2: by Corollary 1 we know that \( K^* < 1 \). In the extreme case, \( K^T = 1 > K^* \).  

Proof of Theorem 1: we know that \( \text{CS}_m > \text{CS} \) if \( \int_{k_b}^{k_p} (p^H - c_b) \theta dF(\theta) + \int_{k_b}^{K} (p^H - c_p) \theta dF(\theta) - m > \int_{k_b}^{k_p} (p^H - c_b) \theta dF(\theta) + \int_{k_b}^{K} (p^H - c_p) \theta dF(\theta) \). By propositions 2 and 3, \( k_b = k_p \). Then \( \text{CS}_m > \text{CS} \iff -m + \int_{k_b}^{K} (p^H - c_p) \theta dF(\theta) > \int_{k_b}^{K} (p^H - c_p) \theta dF(\theta) \). Since we know that in equilibrium \( m = c_k \), then \( \text{CS}_m > \text{CS} \iff p^H - c_p > \int_{k_b}^{K} \frac{1}{K^T} \theta dF(\theta) \).  

Proof of Proposition 4: the price-contingent contract in a market with unregulated capacity, \( \mathcal{L} \), is

\[
\mathcal{L} = \begin{pmatrix}
    p^c_{\text{pr}(\theta < k_b^*)} & p^c_{\text{pr}(k_b^* \leq \theta < K^*)} & p^c_{\text{pr}(K^* \leq \theta)} \\
    \text{pr}(\theta < k_b^*) & \text{pr}(k_b^* \leq \theta < K^*) & \text{pr}(K^* \leq \theta)
\end{pmatrix} = \begin{pmatrix}
    -c_b & -c_p & -p^H \\
    k_b & k_p & 1 - K
\end{pmatrix}
\]

(A.5)

The price-contingent contract in a market with regulated capacity and capacity payments, \( \mathcal{L}^m \), is

\[
\mathcal{L}^m = \begin{pmatrix}
    p^c_{\text{pr}(\theta < k_b^{*, m})} & p^c_{\text{pr}(k_b^{*, m} \leq \theta \leq K^T, m)} \\
    \text{pr}(\theta < k_b^{*, m}) & \text{pr}(k_b^{*, m} \leq \theta \leq K^T, m)
\end{pmatrix} = \begin{pmatrix}
    -c_b & -c_k & -c_p - c_k \\
    k_b & K^T - k_b
\end{pmatrix}
\]

(A.6)

First, we need to show that both price-contingent contracts have the same mean. Let us begin with the mean of \( \mathcal{L} \):

\[
E(\mathcal{L}) = c_b k_b + c_p k_p + p^H (1 - K)
\]

\[
E(\mathcal{L}) = c_b k_b + c_p k_p + p^H k_p - p^H k_b
\]

Next, we calculate the mean for \( \mathcal{L}^m \):

\[
E(\mathcal{L}^m) = (c_b + c_k) k_b + (c_p + c_k)(K^T - k_b)
\]

\[
E(\mathcal{L}^m) = c_b k_b + c_k k_b + c_p + c_k k_b - c_k k_b - c_k k_b
\]
$$E(L_m) = c_bk_b + c_p + c_{kp} - c_p k_b$$

By assumption, $a_p(1 - K) = c_{kp}$; then:

$$E(L_m) = c_bk_b + c_p + a_p - a_p K - c_p k_b$$

$$E(L_m) = c_bk_b + c_p + p^H - c_p - p^H k_b - p^H k_p + c_p k_b + c_p k_p - c_p k_b$$

Thus $E(L) = p^H - a_p k_p - a_b k_b = E(L_m)$.

To continue with the proof, Figure A.1 is helpful. Notice that this Figure captures the two possible scenarios that we may have: $c_p < c_b + c_{kp}$ in subfigure A.1a; and $c_p > c_b + c_{kp}$ in subfigure A.1b. However, as we see below, both scenarios are equivalent, since area of A ($\Delta A$) and area of B ($\Delta B$) are equal.

Figure A.1: Cdf associated to $L$ (red) and cdf associated to $L_m$ (green)

Thus, from subfigure A.1a it follows that if $p^c \in [0, c_p + c_{kp}]$, $\int_0^{p^c} G_L(z)dz > \int_0^{p^c} G_{L_m}(z)dz$. Next, we claim that, for $p^H$ high enough, the area of A plus area of
B, $\Delta A + \Delta B$, is equal to or smaller than the area of $C$, $\Delta C$. Proof: solving for these areas: $\Delta A + \Delta B = c_{kp}(1 - 1 + k_b) + c_{kp}(1 - k_b - 1 + K) \iff \Delta A + \Delta B = c_{kp}K$.

Next, $\Delta C = (p^H - c_p - c_{kp})(1 - K) \iff \Delta C = a_p - c_{kp} - (a_p - c_{kp})K$. Thus, if $c_{kp}K \leq a_p - c_{kp} - (a_p - c_{kp})K \iff p^H \geq \frac{c_{kp}}{1 - K} + c_p$, then $\Delta A + \Delta B \leq \Delta C \square$. Therefore, again using subfigure A.1a as well as the fact that $\Delta cA \leq \Delta \int_k Gc_p \leq \frac{c}{L} \left[p^H - c_p - c_{kp}\right]$, we have that $\Delta cA \leq \frac{c}{L} \left[p^H - c_p - c_{kp}\right]$. Thus, $\forall p^c \in C$, we have that $\int_0^{p^c} G_{L}(z)dz \geq \int_0^{p^c} G_{L}(z)dz$.

Next, from subfigure A.1b it follows that, as long as $p^H \geq \frac{c_{kp}}{1 - K} + c_p$, if $p^c \in [0, c_b + c_{kp})$, $\int_0^{p^c} G_{L}(z)dz \geq \int_0^{p^c} G_{L}(z)dz$. If $p^c \in [c_b + c_{kp}, c_p]$, $\int_0^{p^c} G_{L}(z)dz \geq \int_0^{p^c} G_{L}(z)dz$. And if $p^c \in [c_p, c_p + c_{kp}]$, $\int_0^{p^c} G_{L}(z)dz \geq \int_0^{p^c} G_{L}(z)dz$. Thus, we know that $\Delta A + \Delta B \leq \Delta C$; then it follows that if $p^c \in [c_p + c_{kp}, p^H)$, $\int_0^{p^c} G_{L}(z)dz > \int_0^{p^c} G_{L}(z)dz$, and if $p^c \in [p^H, \infty)$, $\int_0^{p^c} G_{L}(z)dz > \int_0^{p^c} G_{L}(z)dz$. Thus, $\forall p^c \in C$, we have that $\int_0^{p^c} G_{L}(z)dz \geq \int_0^{p^c} G_{L}(z)dz$. \square

**Proof of Proposition 5:** we already know that $E(L) = p^H - a_bk_b - a_pk_p$ and $E(L^m) = (c_b - c_p)k_b + c_p + c_{kp}$. Thus, the variance of $L$ is

$\text{Var}(L) = k_b[c_b - p^H + a_bk_b + a_pk_p] + k_p[c_p - p^H + a_bk_b + a_pk_p] + (1 - k_b - k_p)[p^H - p^H + a_bk_b + a_pk_p]^2$

Rearranging,

$\text{Var}(L) = k_b[p^H^2 + k_pp^H + k_p^2 + 2k_pk_bp^H + 2k_pk_bp^H + k^2p^2 + 2k_pp^2 + k^2p^2] - 2k_pk^2p^H c_b - 2k_pk^2p^H c_p - k^2p^2 + k^2p^2 - k^2p^2 - 2k_pk^2p^H c_b - 2k_pk^2p^H c_p$

$\text{Var}(L) = k_b[a_b^2 + a_p^2 + k_p^2 + k_p^2 + k_p^2 - 2k_pk_p a_b a_p + 2k_p a_b a_p]$

Next, the variance of $L^m$ is

$\text{Var}(L^m) = k_b[c_b + c_{kp} - (c_b - c_p)k_b - c_p - c_{kp}] + (1 - K)[c_p + c_{kp} - (c_b - c_p)k_b - c_p - c_k]$

Rearranging,

$\text{Var}(L^m) = (k_b - k_b^2 - k_b^2)(c_b^2 + c_p^2 - 2c_b c_p)$

By assumption, we know that $p^H \geq c_p$. Assume first that $p^H = c_p$. Then, $\text{Var}(L) = k_b[a_b^2 + a_p^2 + 2k_p a_b a_p] \iff \text{Var}(L) = (k_b - k_b^2)(c_b^2 + c_p^2 - 2c_b c_p)$. In addition, we know that $\text{Var}(L^m) = (k_b - k_b^2 - k_b^2)(c_b^2 + c_p^2 - 2c_b c_p)$. Then, if $p^H = c_p$, then $\text{Var}(L) = (k_b - k_b^2)(c_b^2 + c_p^2 - 2c_b c_p) > (k_b - k_b^2)(c_b^2 + c_p^2 - 2c_b c_p) = \text{Var}(L^m)$.

Next, we show that $\frac{\partial \text{Var}(L)}{\partial p^H} > 0$.

$\text{Var}(L) = k_b[p^H^2 + k_pp^H^2 + k_p^2 + 2k_pk_b p^H + 2k_pk_b p^H + k^2p^2 + 2k_pp^2 + k^2p^2] - 2k_pk^2p^H c_b - 2k_pk^2p^H c_p - k^2p^2 + k^2p^2 - k^2p^2 - 2k_pk^2p^H c_b - 2k_pk^2p^H c_p$. 

\[\text{Var}(L^m) = (k_b - k_b^2)(c_b^2 + c_p^2 - 2c_b c_p) > (k_b - k_b^2)(c_b^2 + c_p^2 - 2c_b c_p) = \text{Var}(L^m).\]
\[
\frac{\partial \text{Var}(\mathcal{L})}{\partial p^H} = 2k_b p^H + 2k_p p^H + 2k_p k_b c_b + 2k_p k_b c_p + 2k_p^2 c_p
\]
\[+ 2k_b^2 c_b - 2k_b c_b - 2k_p c_p - 2k_b^2 p^H - 2k_p^2 p^H - 4k_p k_b p^H =
\]
\[= 2k_b p^H + 2k_p p^H + 2k_p k_b c_b + 2k_p k_b c_p + 2k_b c_p + 2k_p^2 c_b - 2k_b c_b - 2k_p c_p - 2k_b^2 p^H
\]
\[-2k_p^2 p^H - 4k_p k_b p^H =
\]
\[= k_b a_b + k_p a_p - a_b k_b^2 - a_p k_p^2 - 2k_p k_b p^H + k_p k_b c_p + k_p k_b c_b =
\]
\[= a_b(k_b - k_b^2 - k_p k_b) + a_p(k_p - k_p^2 - k_p k_b) = a_b[k_b(1 - k_b - k_p)] + a_p[k_p(1 - k_p - k_p)] > 0
\]
since \((1 - k_p - k_p) > 0\) by Corollary 1.

Moreover, it is straightforward that \(\frac{\partial \text{Var}(\mathcal{L}^m)}{\partial p^H} = 0\). Therefore, for \(p^H > c_p\), then
\(\text{Var}(\mathcal{L}) > \text{Var}(\mathcal{L}^m)\). \(\square\)
Appendix B

Proof of Proposition 6:

1. To show that this is an equilibrium, it is enough to point that if $\theta \leq \left( \frac{n_b-1}{n_b} \right) k_b$, base load generators compete à la Bertrand, leading to the well-known result that the market clearing price is equal to the marginal cost. Since the quantity allocation mechanism is a uniform-price auction, it is enough for the clearing base load generator to bid at marginal cost, while some other base load generators may bid below marginal costs. Second, to show that there is no other equilibrium, assume on the contrary that there exists an equilibrium in which $\forall j$, $p_j^b(\cdot) < c_b$ and/or $p_j^b(\cdot) > c_b$. If $p_j^b(\cdot) < c_b \forall j$, then there is at least one base load generator making negative profit, so there is a strictly profitable deviation for it. If $p_j^b(\cdot) > c_b \forall j$, at least one base generator is strictly better off by undercutting this bid by an arbitrarily $\epsilon > 0$. If $p_j^b(\cdot) < c_b$ for some $j$ and $p_j^{b'}(\cdot) > c_b$ for some $j'$ and the market clearing price is below marginal cost, there is at least one base load generator making negative profit, so there is strictly profitable deviation for it. Finally, $p_j^b(\cdot) < c_b$ for some $j$ and $p_j^{b'}(\cdot) > c_b$ for some $j'$ and the market clearing price is above marginal cost, at least one base generator is strictly better off by undercutting this bid by an arbitrarily $\epsilon > 0$.

2. To show that this is an equilibrium, let us check for strictly profitable deviations for different generators. First, notice that for base load generators $(j,b)$ this cannot be the case, since by deviating they will get the same profit (if $p_j^b(\cdot) \in (0,c_p)$) or less (otherwise). Next, assume that $p_j^b(\cdot) < c_p$. Then $(j',i)$ will obtain less profit (as long as $\epsilon^j$ is small enough $\forall j$). Finally, for peak load generator there is no strictly profitable deviation, since by bidding $p_j^p(\cdot) = c_p$ they still make zero profit and by bidding $p_j^p(\cdot) < c_p$ they make either zero or negative profit.

Finally, to show that this is the only equilibrium, assume on the contrary that there is another equilibrium. If so, the equilibrium market clearing price must be such that $p^* \leq c_p$ (base load generators have excess capacity and are always willing to produce a strictly positive amount). In such equilibrium, the optimal strategy for the peak generators is $p_j^p(\cdot) > p_j^{b'} (\geq p_j^b(\cdot)$ if $p^* = c_p$) where $j'$ is the base load generator that clears the market. Fix the peak load strategies to be
that one. Now, we face different scenarios. First, assume that \( p_{j}^{b}(\cdot) = c_b, \forall j \). In this case, there is a strictly profitable deviation for at least one base load generator \( j' \), which is bidding \( p_{j}'^{b}(\cdot) \in (c_b, c_p) \). Thus, this case cannot be an equilibrium. Next, assume that there is an equilibrium in which \( p_{j}^{b}(\cdot) = a \in (c_b, c_p), \forall j \). In this case, there is a strictly profitable deviation for at least one base load generator \( j', which is bidding \( p_{j}'^{b}(\cdot) = a - \epsilon \) for some arbitrarily small \( \epsilon > 0 \) such that \( \epsilon < \frac{a - c_b}{n_b} \). Thus, this case cannot be an equilibrium. Next, assume that there is an equilibrium in which \( p_{j}^{b}(\cdot) = a \in (c_b, c_p], \forall j \). In this case, there is a strictly profitable deviation for at least one base load generator \( j', which is bidding \( p_{j}'^{b}(\cdot) = a - \epsilon \) for some arbitrarily small \( \epsilon > 0 \) such that \( \epsilon < \frac{a - c_b}{n_b} \). Thus, this case cannot be an equilibrium. Next, assume that there is an equilibrium in which \( p_{j}^{b}(\cdot) = a_{1} \in [c_b, c_p], for some j_1 \) and \( p_{j}^{b}(\cdot) = a_{2} \in [c_b, c_p], for some j_2 \). W.l.o.g assume that \( a_1 < a_2 \). If some \( j_1' \) clears the market, there is a strictly profitable deviation for at least one base load generator \( j_1'' \) which is bidding \( p_{j_1}'^{b}(\cdot) = a_1 + \epsilon < a_2 \). If some \( j_2' \) clears the market, there is a strictly profitable deviation for at least one base load generator \( j_2'' \) which is (slightly) undercutting \( j_2 \) in the same way as in the previous case. The case in which \( a_1 > a_2 \) is symmetric. Finally, the case in which \( p_{j}^{b}(\cdot) > c_p \) where \( j' \) clears the market cannot be an equilibrium, since there is excess capacity and both generators are willing to produce a strictly positive amount.

3. This case is similar to case #1, with the peak load generators competing à la Bertrand.

4. This case is similar to case #2.

5. Assume on the contrary that \( \forall j \) and \( \forall i, p_{i}^{b}(\cdot) < \bar{p} \). Then bidding \( p_{i}^{b}(\cdot) < p_{i}^{b}(\cdot) + \epsilon < \bar{p} \) is a strictly profitable deviation for one generator, for some arbitrarily small \( \epsilon > 0 \).

\[
\begin{align*}
\text{Proof of Proposition 7:} \\
1. \text{From Proposition 6, if } \theta \leq \left( \frac{n_b - 1}{n_b} \right) k_b, \text{ the market clearing price is } c_b. \text{ In the limiting perfect competition case } n_b \to \infty, \text{ which implies } \left( \frac{n_b - 1}{n_b} \right) \to 1. \text{ Thus, if } \theta \leq k_b, \text{ then } p^* = c_b. \text{ It is straightforward that in this case, since the market clearing price is equal to the base load generators marginal cost, no generator makes positive profit.} \\
2. \text{Using a similar reasoning based on Proposition 6, in this case } p^* = c_p. \text{ In this scenario base load generators produce at maximum capacity. Therefore, } \pi_b = c_p k_b - c_b k_b > 0. \text{ Finally, it is straightforward that, since the market clearing price is equal to the peak load generators marginal cost, no peak load generator makes positive profit.} \\
3. \text{The market clearing price is the same as the one in Proposition 6. In this scenario both types of generators produce at maximum capacity. Then, } \pi_b = \bar{p} k_b - c_b k_b > 0 \text{ and } \pi_p = \bar{p} k_p - c_p k_p > 0.
\end{align*}
\]
Proof of Proposition 8: see Proposition 2, Chapter 1.

Proof of Corollary 3: see Corollary 1, Chapter 1.

Proof of Corollary 4: \( k^*_b = F^{-1}\left(1 - \frac{c_{bk} - c_{bp}}{c_p - c_b}\right) \), which does not depend on the market price. However, \( F^{-1}\left(1 - \frac{c_{bp}}{\alpha_p}\right) - F^{-1}\left(1 - \frac{c_{bk} - c_{bp}}{c_p - c_b}\right) \) \( > F^{-1}\left(1 - \frac{c_{bp}}{\alpha_p}\right) - F^{-1}\left(1 - \frac{c_{bh} - c_{bp}}{c_p - c_b}\right) \), which is true since \( \alpha_p > \alpha_{cap} \) where \( \alpha_{cap} = \bar{p} - c_p \) in the presence of a binding cap. It also follows that \( k^*_b + k^*_p > k^*_{b,\text{cap}} + k^*_{p,\text{cap}} \).

Proof of Proposition 9: see Proposition 3, Chapter 1.

Proof of Corollary 5: see Corollary 2, Chapter 1.

Proof of Proposition 10:

1. First, we show that this is an equilibrium. It is straightforward that for the base load there is no strictly profitable deviation: by bidding strictly less than \( c_b \) this generator will incur negative profit. A similar reasoning holds for the dominant peak load generator.

Next, to show that there is no other equilibrium, assume on the contrary that there is an equilibrium in which \( p^b_d(\cdot) < c_b \) and/or \( p^b_d(\cdot) \leq c_b \). If so, the dominant base load generator and/or the dominant peak load generator make negative profits. Therefore, this cannot be an equilibrium. Note: this is true unless \( k^d_b < \theta \), which is the case considered in the footnote. If so, \( \bar{p}_p \in \mathbb{R}^{++} \) is an equilibrium strategy, since the base load fringe clears the market at \( c_b \).

2. This equilibrium (in weakly dominated strategies) is such that (indulging in some abuse of notation) \( p^d_f(\cdot) > c_p \) for all firms in the peak load fringe, \( p^d_b(\cdot) = c_p \) and \( p^d_p(\cdot) > c_p \). Assume on the contrary that \( p^d_f(\cdot) = c_p \). Peak load fringe aggregate profit is still zero, which implies that this is not a strictly profitable deviation for the peak load fringe. On top of that, if \( p^d_p(\cdot) = c_p \) there is no equilibrium, because the optimal deviation for the dominant base load plant is to infinitesimally undercut the bid of the peak load fringe. The same is true if \( p^d_p(\cdot) = c_p \). Finally, notice that \( p^d_p(\cdot) < c_p \) cannot be an equilibrium, since there is a profitable deviation which is \( p^d_p(\cdot) = c_p \). Likewise, \( p^d_p(\cdot) > c_p \) cannot be an equilibrium, since all the market will be served by the peak load fringe.

3. This case is similar to case 1.

4. It is straightforward to see that this is an equilibrium, since the possible deviations for dominant generator will lead to zero profit (i.e. bidding above \( \bar{p} \)) for
them. Next, to show that there is no other equilibrium, assume on the contrary that \( p_i^d(\cdot) > \bar{p} \). In this case, generator \((i,d)\) will not serve the demand. Therefore, it cannot be an equilibrium, since there is a profitable deviation \((p_i^d(\cdot) = \bar{p})\).

\[
\square
\]

**Proof of Corollary 6:**

1. From Proposition 10, since the base load generators clear the market, \( p^s = c_b \). It is straightforward that in this case, since the market clearing price is equal to the base load generators marginal cost, no generator makes positive profit.

2. From Proposition 10, since the peak load generators clear the market, \( p^s = c_p \). In this scenario base load generators produce at maximum capacity. Therefore, \( \pi_b = c_pk_b - c_pk_b > 0 \) and \( \pi_b^d = c_pk_b^d - c_bk_b^d > 0 \). It is straightforward that in this case, since the market clearing price is equal to the base load generators marginal cost, no generator makes positive profit.

3. In this scenario both types of generators produce at maximum capacity. Then, \( \pi_i^f = \bar{p}k_i - c_bk_i > 0 \) and \( \pi_i^d = \bar{p}k_i - c_bk_i > 0 \).

\[
\square
\]

**Proof of Proposition 11:** let us begin with the base load competitive fringe generators; denote \( \pi_b f \) the base load competitive fringe’s profit and \( \mathbb{E} \) the expectation operator. Then, the expected profit for the base load fringe generators is:

\[
\mathbb{E}\pi_b f = \int_0^{q_b^f} 0q_b^f dF(\theta) + \int_{k_b^f}^{k_b^f + k_p^f} (c_p - c_b)q_b^f dF(\theta) + \int_{k_b^f + k_p^f}^{1} (\bar{p} - c_b)q_b^f dF(\theta) - c_b k_b^f \\
\text{s.t. } 0 \leq q_b^f \leq k_b^f
\]

We can simplify the previous expression as follows:

\[
\mathbb{E}\pi_b f = \int_{k_b^f}^{k_b^f + k_p^f} (c_p - c_b)q_b^f dF(\theta) + \int_{k_b^f + k_p^f}^{1} a_bq_b^f dF(\theta) - c_b k_b^f \\
\text{s.t. } 0 \leq q_b^f \leq k_b^f
\]

where \( a_b \equiv \bar{p} - c_b \). Notice that when \( k_b \leq \theta \) the base load produces electricity at maximum capacity. Therefore, in equilibrium, \( q_b^f = k_b^f \) whenever \( k_b \leq \theta \). Thus:

\[
\mathbb{E}\pi_b f = \int_{k_b^f}^{k_b^f + k_p^f} (c_p - c_b)k_b^f dF(\theta) + \int_{k_b^f + k_p^f}^{1} a_bk_b^f dF(\theta) - c_b k_b^f
\]

Solving for the integrals, we arrive at:

\[
\mathbb{E}\pi_b f = (c_p - c_b)k_b^f [F(k_b + k_p^f) - F(k_b^f)] + a_bk_b^f [1 - F(k_b + k_p^f)] - c_b k_b^f
\]
In equilibrium, since we assume free entry and perfect competition, the expected profit must be equal to zero. Thus:

\[(c_p - c_b)k_b^f[F(k_b + k_p^f) - F(k_b^f)] + a_bk_b^f[1 - F(k_b + k_p^f)] - c_{k_b}k_b^f = 0\]

Now let us turn to the peak load competitive fringe generators; denote \(\pi_p\) the peak load competitive fringe’s generator profits. Thus:

\[
\mathbb{E}\pi_p^f = \int_{0}^{k_b^f} 0q_p dF(\theta) + \int_{k_b}^{k_b + k_p^f} 0q_p dF(\theta) + \int_{k_b + k_p^f}^{1} 0q_p dF(\theta) - c_{k_p}k_p^f
\]

s.t. \(0 \leq q_p^f \leq k_p^f\)

We can simplify the previous expression as follows:

\[
\mathbb{E}\pi_p^f = \int_{k_b + k_p^f}^{1} a_p q_p^f dF(\theta) - c_{k_p}k_p^f
\]

s.t. \(0 \leq q_p^f \leq k_p^f\)

where \(a_p \equiv \bar{p} - c_p\). Notice that when \(K \leq \theta\) the peak load produces electricity at maximum capacity. Therefore, in equilibrium, \(q_p^f = k_p^f\) whenever \(k_b + k_p^f \leq \theta\). Thus:

\[
\mathbb{E}\pi_p^f = \int_{k_b + k_p^f}^{1} a_p k_p^f dF(\theta) - c_{k_p}k_p^f
\]

Solving for the integrals, we arrive at:

\[
\mathbb{E}\pi_p^f = a_p k_p^f[1 - F(k_b + k_p^f)] - c_{k_p}k_p^f
\]

In equilibrium, since we assume free entry and perfect competition, the expected profit must be equal to zero. Thus:

\[a_p k_p^f[1 - F(k_b + k_p^f)] - c_{k_p}k_p^f = 0\]

Rearranging, we arrive at:

\[F(k_b + k_p^f) = 1 - \frac{c_{k_p}}{a_p}\]  \hspace{1cm} (B.1)

Plugging equation B.1 into equation B:

\[(c_p - c_b)k_b^f \left[1 - \frac{c_{k_p}}{a_p} - F(k_b^f)\right] + a_bk_b^f \left(\frac{c_{k_p}}{a_p}\right) - c_{k_b}k_b^f = 0\]

Solving for \(k_b^f\) we arrive at:

\[k_b^f = F^{-1}\left(1 - \frac{c_{k_b} - c_{k_p}}{c_p - c_b}\right)\]  \hspace{1cm} (B.2)
Plugging equation B.2 into equation B.1 and solving for $k_p$ we arrive at:

$$k_p^f = F^{-1} \left( 1 - \frac{c_{kp}}{a_p} \right) - F^{-1} \left( 1 - \frac{c_{kp} - c_{kb}}{c_p - c_b} \right) - k_b^d$$  \hspace{1cm} (B.3)

**Proof of Proposition 12:** the base load dominant generator solves

$$k_b^{d,*} = \arg \max_{k_b^d \in [0,1]} (c_p - c_b) \int_{k_b^{f,*}}^{k_b^{d,*} + k_b^d} \theta dF(\theta) + (c_p - c_b)k_b^d \left[ 1 - F(k_b^{f,*} + k_b^d) \right] - k_b^d(c_{kp} - c_{kb})$$  \hspace{1cm} (B.4)

and the peak load generator solves

$$k_p^{d,*} = \arg \max_{k_p^d \in [0,1]} a_p \int_{k_b^{f,*} + k_b^{d,*} + k_p^{d,*}}^{k_b^{f,*} + k_b^{d,*} + k_p^{d,*} + k_p^d} \theta dF(\theta) + a_p k_p^d \left[ 1 - F(k_b^{f,*} + k_b^{d,*} + k_p^{d,*} + k_p^d) \right] - c_{kp} k_p^d$$  \hspace{1cm} (B.5)

Since $F(\cdot)$ is continuous and $[0,1]$ is a closed and bounded interval the extreme value theorem holds, which guarantees the existence of a maximum.

**Proof of Proposition 13:** the proof is similar to the proof of Proposition 11, except that there is a compensation equal to $m$ per unit of capacity built and that $K^T$ denotes total installed capacity.

**Proof of Proposition 14:** the proof is similar to the proof of Proposition 12, except that there is a compensation equal to $m$ per unit of capacity built and that $K^T$ denotes total installed capacity. Thus, the base load dominant generator solves

$$k_b^{d,*} = \arg \max_{k_b^d \in [0,1]} (c_p - c_b) \int_{k_b^{f,*}}^{k_b^{d,*} + k_b^d} \theta dF(\theta) + (c_p - c_b)k_b^d \left[ F(K^T) - F(k_b^{f,*} + k_b^d) \right] + k_b^d(c_{kp} - c_{kb})$$  \hspace{1cm} (B.6)

and the peak load generator solves

$$k_p^{d,*} = \arg \max_{k_p^d \in [0,1]} a_p \int_{k_b^{f,*} + k_b^{d,*} + k_p^{d,*}}^{k_b^{f,*} + k_b^{d,*} + k_p^{d,*} + k_p^d} \theta dF(\theta) +
$$

$$+ a_p k_p^d \left[ F(K^T) - F(k_b^{f,*} + k_b^{d,*} + k_p^{d,*} + k_p^d) \right] - c_{kp} k_p^d + m k_p^d$$  \hspace{1cm} (B.7)

To get the equilibrium compensation mechanism, we know that:

$$K^T = k_b^{*,f,m} + k_b^{*,d,m} + k_p^{*,f,m} + k_p^{*,d,m}$$

Substituting the results obtained in Proposition 13:
\[ K^T = F^{-1} \left[ F(K^T) - \frac{c_{k_p} - m}{a_p} \right] + k^{*,d,m} \]

Rearranging:

\[ m = c_{k_p} + a_p[F(K^T - k^{*,d,m}) - F(K^T)] \]
Appendix C
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min.</th>
<th>Max.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gasoline consumption (t)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unleaded Gasoline 95</td>
<td>7208.007</td>
<td>8161.912</td>
<td>508.928</td>
<td>48408.782</td>
<td>2300</td>
</tr>
<tr>
<td>Unleaded Gasoline 98</td>
<td>579.929</td>
<td>1,107.519</td>
<td>26.559</td>
<td>7977.150</td>
<td>2300</td>
</tr>
<tr>
<td>Diesel A (regular)</td>
<td>33836.523</td>
<td>30568.999</td>
<td>3772.85</td>
<td>187112.182</td>
<td>2300</td>
</tr>
<tr>
<td>Diesel B (agricultural)</td>
<td>6709.114</td>
<td>3956.548</td>
<td>3.925</td>
<td>48105.019</td>
<td>2210</td>
</tr>
<tr>
<td>Real tax-exclusive price (€/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unleaded Gasoline 95</td>
<td>0.719</td>
<td>0.044</td>
<td>0.616</td>
<td>0.912</td>
<td>2300</td>
</tr>
<tr>
<td>Unleaded Gasoline 98</td>
<td>0.786</td>
<td>0.045</td>
<td>0.691</td>
<td>0.989</td>
<td>2300</td>
</tr>
<tr>
<td>Diesel A (regular)</td>
<td>0.759</td>
<td>0.037</td>
<td>0.653</td>
<td>0.915</td>
<td>2300</td>
</tr>
<tr>
<td>Diesel B (agricultural)</td>
<td>0.763</td>
<td>0.036</td>
<td>0.627</td>
<td>0.859</td>
<td>2208</td>
</tr>
<tr>
<td>Real tax (€/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unleaded Gasoline 95</td>
<td>0.668</td>
<td>0.093</td>
<td>0.218</td>
<td>0.738</td>
<td>2300</td>
</tr>
<tr>
<td>Unleaded Gasoline 98</td>
<td>0.716</td>
<td>0.103</td>
<td>0.218</td>
<td>0.789</td>
<td>2300</td>
</tr>
<tr>
<td>Diesel A (regular)</td>
<td>0.564</td>
<td>0.087</td>
<td>0.103</td>
<td>0.633</td>
<td>2300</td>
</tr>
<tr>
<td>Diesel B (agricultural)</td>
<td>0.254</td>
<td>0.016</td>
<td>0.213</td>
<td>0.285</td>
<td>2208</td>
</tr>
<tr>
<td># vehicles using unleaded fuels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vans &amp; Trucks</td>
<td>10390.962</td>
<td>13048.032</td>
<td>26</td>
<td>73207</td>
<td>2250</td>
</tr>
<tr>
<td>Buses</td>
<td>15.606</td>
<td>25.131</td>
<td>1</td>
<td>168</td>
<td>2161</td>
</tr>
<tr>
<td>Cars</td>
<td>193037.243</td>
<td>264995.111</td>
<td>82</td>
<td>1474019</td>
<td>2250</td>
</tr>
<tr>
<td>Motorcycles</td>
<td>52897.375</td>
<td>83642.042</td>
<td>72</td>
<td>544932</td>
<td>2250</td>
</tr>
<tr>
<td>Tractors</td>
<td>30.941</td>
<td>33.969</td>
<td>2</td>
<td>238</td>
<td>2250</td>
</tr>
<tr>
<td>Other vehicles</td>
<td>2367.825</td>
<td>2359.195</td>
<td>11</td>
<td>14216</td>
<td>2250</td>
</tr>
<tr>
<td># vehicles using diesel fuels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vans &amp; Trucks</td>
<td>80778.41</td>
<td>89403.415</td>
<td>147</td>
<td>559952</td>
<td>2250</td>
</tr>
<tr>
<td>Buses</td>
<td>1122.624</td>
<td>1509.159</td>
<td>14</td>
<td>10484</td>
<td>2250</td>
</tr>
<tr>
<td>Cars</td>
<td>223544.736</td>
<td>300226.251</td>
<td>49</td>
<td>1924655</td>
<td>2250</td>
</tr>
<tr>
<td>Motorcycles</td>
<td>46.264</td>
<td>55.976</td>
<td>1</td>
<td>324</td>
<td>2250</td>
</tr>
<tr>
<td>Tractors</td>
<td>3446.931</td>
<td>3370.651</td>
<td>2</td>
<td>17832</td>
<td>2250</td>
</tr>
<tr>
<td>Other vehicles</td>
<td>5710.279</td>
<td>4912.174</td>
<td>7</td>
<td>29398</td>
<td>2250</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing (real €)</td>
<td>96640.476</td>
<td>74921.036</td>
<td>21769.69</td>
<td>2443541.039</td>
<td>2300</td>
</tr>
<tr>
<td>Credits (real €)</td>
<td>146715.048</td>
<td>2870382.143</td>
<td>18831.823</td>
<td>137534733.033</td>
<td>2300</td>
</tr>
<tr>
<td>Agric. unemploy. (#)</td>
<td>3611.99</td>
<td>3224.225</td>
<td>157</td>
<td>17304</td>
<td>2300</td>
</tr>
<tr>
<td>Brent (€/l)</td>
<td>0.514</td>
<td>0.031</td>
<td>0.434</td>
<td>0.597</td>
<td>2300</td>
</tr>
</tbody>
</table>
Table C.2: Summary statistics for (nominal) tax data (by Autonomous Community)

<table>
<thead>
<tr>
<th>Auton. Com.</th>
<th>Statistic</th>
<th>Diesel A</th>
<th>Diesel B</th>
<th>Unleaded 95</th>
<th>Unleaded 98</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalucia</td>
<td>Mean</td>
<td>0.59</td>
<td>0.25</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.54</td>
<td>0.22</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Aragon</td>
<td>Mean</td>
<td>0.54</td>
<td>0.25</td>
<td>0.65</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.51</td>
<td>0.21</td>
<td>0.61</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.57</td>
<td>0.28</td>
<td>0.68</td>
<td>0.73</td>
</tr>
<tr>
<td>Asturias</td>
<td>Mean</td>
<td>0.58</td>
<td>0.26</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.54</td>
<td>0.23</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.62</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Balearic Islands</td>
<td>Mean</td>
<td>0.59</td>
<td>0.26</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.05</td>
<td>0.02</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.51</td>
<td>0.22</td>
<td>0.62</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Canary Islands</td>
<td>Mean</td>
<td>0.17</td>
<td>-</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.06</td>
<td>-</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.10</td>
<td>-</td>
<td>0.22</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.22</td>
<td>-</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>Cantabria</td>
<td>Mean</td>
<td>0.58</td>
<td>0.25</td>
<td>0.68</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.04</td>
<td>0.02</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.51</td>
<td>0.22</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Castilla-La Mancha</td>
<td>Mean</td>
<td>0.59</td>
<td>0.25</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.54</td>
<td>0.22</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Castilla and Leon</td>
<td>Mean</td>
<td>0.58</td>
<td>0.25</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.04</td>
<td>0.02</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.51</td>
<td>0.22</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Catalonia</td>
<td>Mean</td>
<td>0.59</td>
<td>0.26</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.54</td>
<td>0.23</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.29</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Ceuta</td>
<td>Mean</td>
<td>0.16</td>
<td>-</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.16</td>
<td>-</td>
<td>0.20</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.16</td>
<td>-</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>Extremadura</td>
<td>Mean</td>
<td>0.59</td>
<td>0.25</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.54</td>
<td>0.22</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
<tr>
<td>Galicia</td>
<td>Mean</td>
<td>0.57</td>
<td>0.25</td>
<td>0.68</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.53</td>
<td>0.22</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.62</td>
<td>0.28</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td>Madrid</td>
<td>Mean</td>
<td>0.57</td>
<td>0.26</td>
<td>0.67</td>
<td>0.72</td>
</tr>
</tbody>
</table>

*Continued on next page*
<table>
<thead>
<tr>
<th>Auton. Com.</th>
<th>Statistic</th>
<th>Diesel A</th>
<th>Diesel B</th>
<th>Unleaded 95</th>
<th>Unleaded 98</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Melilla</td>
<td>Min</td>
<td>0.53</td>
<td>0.23</td>
<td>0.64</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.59</td>
<td>0.28</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>0.10</td>
<td>0.00</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.00</td>
<td>-</td>
<td>0.00</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.09</td>
<td>0.00</td>
<td>0.16</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.10</td>
<td>0.00</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>Murcia</td>
<td>Mean</td>
<td>0.58</td>
<td>0.26</td>
<td>0.69</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.04</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.52</td>
<td>0.23</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.62</td>
<td>0.28</td>
<td>0.73</td>
<td>0.78</td>
</tr>
<tr>
<td>Navarra</td>
<td>Mean</td>
<td>0.55</td>
<td>0.25</td>
<td>0.66</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.03</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.51</td>
<td>0.21</td>
<td>0.61</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.60</td>
<td>0.28</td>
<td>0.71</td>
<td>0.76</td>
</tr>
<tr>
<td>Basque Country</td>
<td>Mean</td>
<td>0.55</td>
<td>0.26</td>
<td>0.65</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.51</td>
<td>0.22</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.57</td>
<td>0.28</td>
<td>0.68</td>
<td>0.73</td>
</tr>
<tr>
<td>La Rioja</td>
<td>Mean</td>
<td>0.55</td>
<td>0.25</td>
<td>0.65</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.51</td>
<td>0.22</td>
<td>0.62</td>
<td>0.66</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.57</td>
<td>0.28</td>
<td>0.68</td>
<td>0.73</td>
</tr>
<tr>
<td>Valencia</td>
<td>Mean</td>
<td>0.59</td>
<td>0.25</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>Std. dev.</td>
<td>0.04</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Min</td>
<td>0.53</td>
<td>0.22</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>Max</td>
<td>0.63</td>
<td>0.28</td>
<td>0.74</td>
<td>0.79</td>
</tr>
</tbody>
</table>

Source: Comisión Nacional de los Mercados y de la Competencia (CNMC), Jan’11-Oct’14.
Appendix D

Proof of Claim 1: since the transportation cost exhibits the same quadratic growth for all consumers along the interval, solving $\hat{x}$ in equation 5.6 is similar to calculate the point $\hat{x}$ in between $x_1$ and $x_2$ such that the areas to the left of $\hat{x}$ until $x_1$ and to the right of $\hat{x}$ until $x_2$ are equal. I.e.

$$\int_{x_1}^{\hat{x}} f(x)dx = \int_{\hat{x}}^{x_2} f(x)dx$$

(D.1)

which implies that $\hat{x}$ is determined by

$$F(\hat{x}) = \frac{F(x_2) + F(x_1)}{2}$$

(D.2)

To show that $\hat{x}$ is unique, a.f.s.o.c. that there exists $\hat{x}'$ and $\hat{x}''$ that satisfy equation D.2. W.l.o.g. assume $\hat{x}' > \hat{x}''$. By strictly monotonicity of $F(\cdot)$, then $F(\hat{x}') > F(\hat{x}'')$. Thus, $F(\hat{x}') = \frac{F(x_2) + F(x_1)}{2} > F(\hat{x}'')$, which is a contradiction to the fact that both $\hat{x}'$ and $\hat{x}''$ satisfy equation D.2. The proof for the case $\hat{x}' < \hat{x}''$ is similar.

$\square$
Appendix E

We propose some examples to see how firms’ locations, the distribution of consumers and transportation costs (all of them exogenously given) are critical variables that determine the equilibrium prices.

We propose four different scenarios. First, we assume a baseline scenario, characterized by a symmetric distribution of consumers and symmetric locations. The transportation cost is assumed to be $\tau = 25$. We deviate from this case in the second, third and fourth scenarios, in which we assume instead different $\tau$, asymmetric distribution of consumers and asymmetric firms’ locations respectively. For the sake of simplicity, we impose the assumption that $c(\theta) = 0$ for $i \in \{1, 2\}$ for all $\theta \in \Theta$. Moreover, we normalize the interval of the consumers to $[0, 1]$. Finally, we assume that consumers are distributed according to a beta distribution.

Taking into account that most of the commonly used distributions –including the beta distribution– do not have a close form representation of their cdf, equations 5.4 and 5.5 are usually not easily tractable. Therefore, we use the collocation method proposed in Fackler and Miranda (2004) to approximate the best-response function for each firm given the other firm’s best-response function.

For that purpose, the best response function, i.e. $p_i^*(\cdot)$, is approximated using a linear combination of $n$ basis functions such that:

$$\hat{p}_i^*(\cdot) = \sum_{j=1}^{n} c_j \phi_j(p_{-i})$$  \hspace{1cm} (E.1)

for the $n$ fixed coefficients $n_j$ (called collocation nodes) at which the approximant is imposed to satisfy the functional form. Thus, by defining function $g : \mathbb{R}^2 \rightarrow \mathbb{R}$, such that

$$g\left(p_{-i}, \sum_{j=1}^{n} c_j \phi_j(p_{-i})\right) = 0 \quad \text{for} \quad j \in \{1, 2, \ldots, n\}$$  \hspace{1cm} (E.2)

we can reduce the previous problem to a system of $n$ non-linear equations with $n$ unknowns ($c_j$) that can be solved using standard rootfinding techniques, such as Boydren’s method or Newton’s method. In our case, we use the former one. Moreover, to approximate the best response functions, we use a 20-degree Chebychev polynomial as approximant (basis function). The results for different beta shape parameters and locations are displayed in Figure E.1.
Figure E.1: Equilibrium prices for different firms’ locations, distributions of consumers and transportation costs

(a) Base case: symmetric distribution and locations ($\tau = 25$)

(b) Symmetric distribution and locations, lower transportation cost ($\tau = 24$)

(c) Demand advantage for firm 1: left skewed distribution, symmetric location ($\tau = 25$)

(d) Location advantage for firm 1, symmetric distribution ($\tau = 25$)
First, we include the baseline case in subfigure E.1a. We assume a symmetric beta distribution, with parameters $\alpha = \beta = 2$, firms’ symmetric locations—in particular $x_1 = \frac{1}{3}$ and $x_2 = \frac{2}{3}$—, and transportation cost $\tau = 25$. Not surprisingly, the unique equilibrium price for this case is symmetric: both firms’ prices are approximately 5.6. In subfigure E.1b we modify the baseline case by decreasing the transportation cost from $\tau = 25$ to $\tau = 24$. As a result, since it is less costly for consumers to travel to the firms, equilibrium prices decrease to roughly 5.4—in the extreme case in which $\tau = 0$, firms simply reduce their prices to the highest marginal cost (Bertrand competition).

Next, in subfigure E.1c we have considered a slightly left-skewed distribution of consumers with $\tau = 25$ (as in the base case). Therefore, firm 1 is now closer to the point with highest concentration of consumers, giving an advantage to it. On the other hand, since the bulk of consumers is farther from firm 2, this firm needs to lower its price to attract more consumers. More precisely, firm 1’s price is approximately 5.6, while firm 2’s price is close to 5.1. Finally, subfigure E.1d contains the case of symmetric demand distribution but asymmetric locations with $\tau = 25$. In particular, we assume that firm 1 is now closer to the center, giving again an advantage to it over firm 2. In this case, notice that firm 1’s optimal price is approximately 4.6, while firm 2’s price is 4.4. Notice that it also make sense to see a reduction in the equilibrium prices when both firms are closer to each other—in the extreme case in which both firms are located in the same point (i.e. $x_1 = x_2$), firms simply compete à la Bertrand.

These graphs give us an idea how important are firms’ location, the distribution of consumers and the transportation cost to determine the optimal strategies.
Appendix F

We conduct a simulation study to illustrate how we partially identify $T$ points of the distribution and how we recover a close enough density function if we have sufficient unique points. For that purpose, we consider the following: first, firms’ marginal costs at each period are drawn from a uniform distribution between 0.7 and 1.2. I.e. $c(\theta^t) \sim U[0.7, 1.2]$ for all $t$. Second, locations are such that $x_1 = \frac{1}{3}$ and $x_2 = \frac{2}{3}$, with the interval normalized to $[0, 1]$. Third, the (known) consumers’ transportation cost is $\tau = 1.5$. Next, we draw 2,000 costs parameters to obtain another 2,000 firms’ equilibrium prices for each pair of variable costs drawn. Using them, we are able to identify (at most) 2,000 Cartesian points of the distribution $f(\cdot)$ plugging all the observed relevant parameters in equation 5.12.

Let $\lambda_t$ and $\theta_t$ be an element in the range and in the domain of $f(\cdot)$ respectively. Thus, having up to observation $T$, call $(\lambda_1, \theta_1), \ldots, (\lambda_T, \theta_T)$ the available sample of points of the distribution. To get an approximation of the density function based upon these identified points, we use the local polynomial regression between $[x, \bar{x}] = [0, 1]$.

A detailed explanation of such estimation method for the case in which $\theta$ is unknown is included in section 3.

We comment one remark on the local polynomial regression. As Fan and Gijbels (1995) point out, choosing the order of polynomial does not come for free; in fact, they show that there is a tradeoff between choosing a higher order for the polynomial, which reduces bias at the cost of losing efficiency, and choosing a lower order, which increases variability in the estimation at the potential cost of increasing the bias. In this particular case, we have considered both a second order polynomial approximation and a fourth order polynomial approximation. The second order polynomial one is reasonable if we believe that the true distribution is a (strictly) concave function. However, the fourth order polynomial is able to capture further changes in the slope (peaks and valleys) that log-concave (but not concave) functions may present.

To show how our strategy performs for different distributions of consumers, we have considered several scenarios. First, we assume that the true distribution of consumers over the interval $[0, 1]$ follows a beta distribution. We have considered different shape parameters that allows us to check differences in skewness and kurtosis. The results for the simulations with a beta distribution are included in Figure F.1. We have considered a symmetric beta distribution ($\alpha = 2, \beta = 2$), which is plotted in the middle; a relatively high-skewed distribution whose center is towards the left ($\alpha = 2.5, \beta = 5.5$), plotted in the top of the figure; and a relatively low-skewed distribution
whose center is towards the right ($\alpha = 4.5, \beta = 3$), included in the bottom.

Figure F.1: True and fitted distribution of the demand (true follows a beta)

We use a bandwidth $b = 0.14$ for the approximation of order 2 and $b = 0.19$ for the approximation of order 4. A Gaussian kernel was considered.

As we can see, the methodology works pretty well in the symmetric case (middle). Both the quadratic and the fourth order polynomials are able to capture the whole shape of the true distribution. Regarding the skewed distribution, the quadratic approximation still does a good job in capturing the center of the distribution. However, the approximation is not as accurate for the opposite side where the distribution is skewed. In these cases, the fourth order approximation does a better job in capturing the curvature of the peaks as well as the curvature in the valleys of the distribution.

Next, we check the estimated distribution considering that the true one is a quadratic function of the sort $f(x) = ax^2 + bx + c$, where $f(x) = 0$ if $x \not\in [0,1]$ and $\int_x f(x)dx = 1$. Again, we have considered different shapes by assuming differ-
ent true values for $a$, $b$ and $c$. The plots of the estimated and real distributions are included in Figure F.2.

Figure F.2: True and fitted distribution of the demand (true follows a quadratic)

![Graphs showing estimated vs. real distributions for different polynomials](image)

We use a bandwidth $b = 0.13$ for the approximation of order 2 and $b = 0.14$ for the approximation of order 4. A Gaussian kernel was considered.

In this case, both the estimated quadratic and the fourth order polynomials approximate very accurately the real distribution. Only a slight deviation is observed for the fourth order polynomial case, which deviates a little bit from the true distribution for values close to 1 for the left-skewed distribution (top right), and for values close to 0 for the right-skewed distribution (bottom right).
Appendix G

Residual functions for Figure E.1

Figure G.1: Base case: symmetric distribution and locations ($\tau = 25$)

Figure G.2: Symmetric distribution and locations, lower transportation cost ($\tau = 24$)
Figure G.3: Demand advantage for firm 1: left skewed distribution, symmetric location ($\tau = 25$)

Figure G.4: Location advantage for firm 1, symmetric distribution ($\tau = 25$)