## Title
Essays in Industrial Organization: Entry in Multi-Object Auctions and Freemium Packages in Two-sided Markets



## Author
Zaterka Giroldo, Renato

## Publication Date
2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

# Essays in Industrial Organization
# Entry in Multi-Object Auctions and Freemium Packages in Two-sided Markets

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Economics

by

Renato Zaterka Giroldo

2019

ABSTRACT OF THE DISSERTATION

Essays in Industrial Organization

Entry in Multi-Object Auctions and Freemium Packages in Two-sided Markets

by

Renato Zaterka Giroldo

Doctor of Philosophy in Economics

University of California, Los Angeles, 2019

Professor John William Asker, Chair

In the first two chapters of this dissertation, I study the design of multi-object auctions. Using a large data set from the Brazilian public procurement sector, I show evidence that entry is costly and that the mix of products being auctioned off is a first-order effect to understand firm participation.

In the first chapter, I find evidence that the data is consistent with a theory of selection. The average entrant has a higher product match with the session, and they are closer to auction locations. Distance affects entry decisions negatively: a 1 unit (100km) increase in the distance to the auction location lowers the odds ratio for entry 0.91 times. At the same time, an additional auction in the set of potential auctions of a firm increases the odds for entry 1.62 times.

In terms of variable costs, a 1% increase in the distance to the auction location increases bid by 0.4% to 3.3%. There are also gains of scale in terms of the size of the contracts: a 1% increase in the contract quantity for a given product increases bids by 0.64% to 0.76%. The main force responsible for lowering procurement costs is the presence of additional bidders. I find that an extra bidder can lower costs between 21.2% and 32.4%. These results motivate and feed into the structural model presented in chapter 2.

In the second chapter, I continue to analyze this market with the focus on estimating entry costs and answering policy questions. To do so, I build a novel model of endogenous

entry in multi-object auction sessions that allows me to disentangle two forces that affect entry decisions: entry costs, and the menu of items of a given session. The model has two stages. In the first stage, firms decide whether to enter an auction session and pay a fixed cost after observing an imperfect signal of their true cost. In the second stage, both the items for which they can bid and their costs are realized, and the auction takes place. I focus the analysis on type symmetric equilibria, where bidders of the same type follow the same entry strategy. In equilibrium, marginal bidders make zero profits. This condition allows me to link the unobserved entry costs to the observed bid behavior of entrants.

Having derived the equilibrium of the model, I estimate model fundamentals and turn to policy questions. The estimates provide evidence that entry is more attractive to local firms. I find that their cost distribution stochastically dominates the one from non-local firms. Moreover, conditioned on the number of items a firm can participate in, non-local firms face between 2.9% to 7.1% higher entry costs than local firms.

I focus on two counterfactual simulations. In the fully efficient scenario, where firms do not incur any entry costs, I find that procurement costs would be lowered by 22.5% to 40.1%. These are bounds on the maximum cost savings and also quantifies the degree of inefficiency present in this market. The second counterfactual is a partially efficient scenario, where non-local firms face the same entry costs as local firms. This analysis focuses on a selected equilibrium where firms enter the sessions sequentially. Firms are sorted according to a lexicographic order which is determined by the strength of their signal, number of items, and firm type (non-local/local). I find that procurement costs would be lowered by 2.8% to 2.9%. Thus, on this type of equilibrium and by holding on-site auctions, the government indirectly sacrificed some efficiency to the benefit of local firms.

In the third chapter, I study the pricing of platforms that offer consumers the choice between a free package, in which consumers are exposed to advertising, and a premium package, in which they pay to not be exposed to advertisements. I characterize its profit-maximizing and Pigouvian pricing, which allows me to analyze the degree to which the platform incorporates consumers' distaste for advertising in its pricing scheme, as well as the trade-offs that emerge between the free and paid packages. The results contribute to the discussion of consumers' overexposure to advertising when platforms behave as a social planner and maximize their value.

The dissertation of Renato Zaterka Giroldo is approved.

Rosa Liliana Matzkin

Brett William Hollenbeck

Simon Adrian Board

John William Asker, Committee Chair

University of California, Los Angeles

2019

*To my mother Beatriz and my father José Maria*
*for their love, care, and support throughout this journey.*

TABLE OF CONTENTS

**1 Menu Choice and Market Structure:**
**Selection in the Brazilian On-site Procurement Sector . . . . . . . . . . . . . 1**

1.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

1.2 Literature review . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4

1.3 Data description . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6

1.3.1 Letting process: on-site sections . . . . . . . . . . . . . . . . . . . 7

1.3.2 Descriptive Statistics . . . . . . . . . . . . . . . . . . . . . . . . . 8

1.4 Stylized Facts . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

1.4.1 Two-step design . . . . . . . . . . . . . . . . . . . . . . . . . . . . 10

1.4.2 Evidence of selection . . . . . . . . . . . . . . . . . . . . . . . . . 11

1.4.3 Firm size distribution . . . . . . . . . . . . . . . . . . . . . . . . . 12

1.5 Drivers of entry and procurement cost . . . . . . . . . . . . . . . . . . . 13

1.6 Conclusion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16

1.7 Appendix I: Tables . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

1.8 Appendix II: Reduced-form results . . . . . . . . . . . . . . . . . . . . . 21

1.9 Appendix III: Figures . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25

**2 Menu Choice and Market Structure:**
**An empirical exercise on multi-object auctions . . . . . . . . . . . . . . . . . 37**

2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 37

2.2 Structural Model . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39

2.2.1 Fundamentals . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39

2.2.2 Signal structure . . . . . . . . . . . . . . . . . . . . . . . . . . . . 40

2.2.3 Timing . . . . . . . . . . 40

2.2.4 Assumptions . . . . . . . . . . 40

2.2.5 Equilibrium . . . . . . . . . . 41

2.2.6 Type space . . . . . . . . . . 43

2.3 Identification . . . . . . . . . . 44

2.4 Estimation Procedure . . . . . . . . . . 47

2.5 Empirical Results . . . . . . . . . . 50

2.6 Conclusion . . . . . . . . . . 53

2.7 Appendix I: Proofs . . . . . . . . . . 54

2.7.1 Proof of Proposition 2.2.1 . . . . . . . . . . 54

2.8 Appendix II: Figures . . . . . . . . . . 58

2.8.1 Bounds on order statistics . . . . . . . . . . 58

2.8.2 Bounds on cost distribution . . . . . . . . . . 60

2.8.3 Shocks on entry . . . . . . . . . . 62

2.9 Appendix III: Tables . . . . . . . . . . 63

2.9.1 Profits . . . . . . . . . . 63

2.9.2 Entry Costs . . . . . . . . . . 63

**3 Freemium Packages in Two-Sided Markets . . . . . . . . . . 64**

3.1 Introduction . . . . . . . . . . 64

3.2 Literature Review . . . . . . . . . . 65

3.3 Monopolistic Platform . . . . . . . . . . 67

3.3.1 Equilibrium Prices . . . . . . . . . . 71

3.3.2 Comparison with Social Planner . . . . . . . . . . 73

3.4 Conclusion . . . . . . . . . . 76

3.5 Appendix I: Identification . . . . . . . . . . . . . . . . . . . . . . . . . . 78

3.6 Appendix II: Simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . 83

3.7 Appendix III: Platform Competition . . . . . . . . . . . . . . . . . . . . . 88

3.8 Appendix IV: Proofs . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 91

3.8.1 Proof of Proposition 3.3.1 . . . . . . . . . . . . . . . . . . . . . . . . 91

3.8.2 Proof of Proposition 3.3.2 . . . . . . . . . . . . . . . . . . . . . . . . 93

3.8.3 Proof of Proposition 3.7.1 . . . . . . . . . . . . . . . . . . . . . . . . 95

3.8.4 Proof of Proposition 3.5.1 . . . . . . . . . . . . . . . . . . . . . . . . 100

3.8.5 Proof of Proposition 3.5.2 . . . . . . . . . . . . . . . . . . . . . . . . 101

3.8.6 Proof of Proposition 3.5.3 . . . . . . . . . . . . . . . . . . . . . . . . 104

LIST OF FIGURES

1.1 Item 2 description (paper towel) of session 1/2005 held by the National School of Public Administration in Brasilia, buyer code 114702. Session was held at 03/14/2005. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25

1.2 Item 2 description (paper towel) of session 1/2005 held by the National School of Public Administration in Brazilia, buyer code 114702, on 03/14/2005. The bid history is separated by the proposals phase "Propostas" and bid phase "Lances". The variable "CNPJ/CPF" identifies a firm. Firms with "D" next to their proposal were disqualified. Firms with "*" next to their proposal moved to the open descending phase which is identified by "Lances". Dropout points are market with "Desistiu". . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 25

1.3 Monthly number of sessions. Sample consists of all on-site sessions between 2003 and 2018 regarding federal purchases of services and off-the-shelf goods, procured through open descending price format ("pregao"). On June 2005, the government regulated online auctions making the electronic format the preferred procedure for the federal acquisition of this type of contract. Law 5.450-05, Art. 4. . . . . 26

1.4 Monthly number of sessions. Sample consists of all online sessions between 2003 and 2018 regarding federal purchases of services and off-the-shelf goods, procured through open descending price format ("pregao"). On June 2005, the government regulated online auctions making the electronic format the preferred procedure for the federal acquisition of this type of contract. Law 5.450-05, Art. 4. . . . . 26

1.5 Total number of sessions by buyer location between 2003 and 2006. A buyer is an administrative unit of general service (UASG). . . . . . . . . . . . . . . . . . 27

1.6 Total number of sessions by state in the years of 2003, 2004, 2005, and 2006. . . 27

1.7 Seller location. Small firms in red and big firms in blue. . . . . . . . . . . . . . . 28

1.8 Histogram of distances from winners to auction location. Small firms on the left and big firms on the right. Sample consists of off-the-shelf goods. . . . . . . . . 28

1.9 Potential participation in a session vs actual participation. Potential participation corresponds to the number of auctions whose product classes belong to the firm portfolio. . . . . . . . . . 29

1.10 Contract size distribution. Histogram truncated at $10,000. . . . . . . . . . . 29

1.11 Procurement costs distribution allocated to small and big firms. Sample restricted to off-the-shelf contracts with value between $1,000 and $10,000. . . . . . . . . . 30

1.12 Equally-weighted share of contracts by distance to auction location. Sample restricted to 'off the shelf' sessions. . . . . . . . . . 31

1.13 Equally-weighted share of 'services' contracts by distance to auction location. Sample restricted to 'services' sessions. . . . . . . . . . 31

1.14 Value-weighted share of contracts by distance to auction location. Sample restricted to 'off the shelf' sessions. . . . . . . . . . 32

1.15 Value-weighted share of contracts by distance to auction location. Sample restricted to 'services' sessions. . . . . . . . . . 32

1.16 Equally-weighted share of contracts by firm size. Sample restricted to 'off the shelf' sessions. . . . . . . . . . 33

1.17 Equally-weighted share of contracts by firm size. Sample restricted to 'services' sessions. . . . . . . . . . 33

1.18 Value-weighted share of contracts by firm size. Sample restricted to 'off the shelf' sessions. . . . . . . . . . 34

1.19 Value-weighted share of contracts by firm size. Sample restricted to 'services' sessions. . . . . . . . . . 34

1.20 Equally-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'off the shelf' sessions. . . . . . . . . . 35

1.21 Equally-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'services' sessions. . . . . . . . . . 35

1.22 Value-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'off the shelf' sessions. . . . . . . . . . . . . . . . . . . . . . . 36

1.23 Value-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'services' sessions. . . . . . . . . . . . . . . . . . . . . . . . . 36

2.1 Bounds on conditional order statistics for firm type *far*. Estimated with auctions where N players entered and achieved the z order statistic with their bids. They were constructed using the residuals from section 2.4.1. Upper bounds in blue and lower bounds in red. Figure presents results for the firm type *far* and $\bar{s} = 1$. Here, the firm type is *far* if it is more than 300km away from the session. . . . . 58

2.2 Bounds on conditional order statistics for both firm types *close* and *far*. Estimated with auctions where N=3 players entered and achieved the z=2 order statistic with their bids. They were constructed using the residuals from section 2.4.1. Firm type *close* in blue and *far* type in red. Figure presents results for $\bar{s} = 1$. Here, the firm type is *far* if it is more than 300km away from the session. 59

2.3 Bounds on conditional cost distribution for firm type *far*. Estimated with auctions where N players entered and achieved the z order statistic with their bids. Constructed using bounds on order statistics and the mapping between those and cost distributions (equation 2.6). Upper bounds in blue and lower bounds in red. Figure presents results for the firm type *far* and $\bar{s} = 1$. Here, firms are far if they are more than 300km away from the session. . . . . . . . . . . . . . . . . . . . . 60

2.4 Smoothed estimators of the minimum upper bound and maximum lower bound for the conditional cost distribution. Firm type *close* in blue and *far* in red. Figure presents results estimated with auctions up to N=3, and $\bar{s} = 1$. Here, firms are far if they are more than 300km away from the session. . . . . . . . . . 61

2.5 Scatter plot at the session level of the number of auctions in the firms' feasible set versus the fraction of auctions participated in. Firms feasible sets are constructed as the intersection between the firms portfolio and the menu of items auctioned at the session. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 62

3.1 Platform Profits: Two sided vs One sided . . . . . . . . . . . . . . . . . . . . . . . . 85
3.2 Consumer Prices: Two sided vs One sided . . . . . . . . . . . . . . . . . . . . . . . . 86
3.3 Figure 3.1 and Figure 3.2: views from above . . . . . . . . . . . . . . . . . . . . . . 87
3.4 Figure 3.1 and Figure 3.2 overlaid . . . . . . . . . . . . . . . . . . . . . . . . . . . . 87

LIST OF TABLES

1.1 List of variables. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17

1.2 Statistics - all Sessions. Final sample consists of all sessions with value between \$1,000 and \$10,000,000 and individual auctions with value above \$100. . . . . . 17

1.3 Top 5 product classes for all on-site sessions. A product class is defined by a four digit code where the first two numbers represent the product group it belongs to. 18

1.4 Top 5 Buyers. A buyer is an administrative unit of general service (UASG). . . 18

1.5 Overview of auction outcomes. Values adjusted for inflation. Base is Jan/2010. Sample consists of all sessions with value between \$1,000 and \$1,000,000 and auctions with value above \$100. . . . . . . . . . . . . . . . . . . . . . . . . . . 19

1.6 Rank inversion between first and second phases of the auctions. Sample consists of all 21,602 off-the-shelf auctions with more than three participants. . . . . . . 20

1.7 Rank inversion between first and second phases of the auctions. Sample consists of all 27,403 off-the-shelf auctions with less than or equal to three participants. . 20

1.8 Outcomes from entrants compared to firms that did not participate. All off-the-shelf sessions in the sample. Distance in multiples of 100km. . . . . . . . . . . . 20

1.9 Conditional fixed effects logistic regression of entry decisions on session characteristics. Sample consists of potential participants for off-the-shelf sessions. Total number of observations in the sample is 3,116,600. Small is a dummy variable with value equal to one if the firm is *small*. . . . . . . . . . . . . . . . . . . . . . 21

1.10 Heckman selection model. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 22

1.11 Two-stage least squares regression of winning bids on observables. Distance in 100km units. The variable "Log Bid" corresponds to the procurement cost of the contract. Values in $ adjusted for inflation. Base is Jan/2010. Specification includes product, supply type, and location fixed effects as well as clusters for location and product. The variable 'Participation' is instrumented with the variable 'portfolio match' which uses 6 digit product codes. Small is a dummy variable with value equal to one if the firm is *small*. . . . . . . . . . . . . . . . . . 24

2.1 Bounds on mean session profits by entry decisions and firm type. . . . . . . . . 63

2.2 Entry cost estimates. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 63

## ACKNOWLEDGMENTS

I would like to thank Professors John Asker, Simon Board, Rosa Matzkin, Brett Hollenbeck, and Martin Hackmann for the valuable lessons, talks, and meetings that helped shape this work; and, more importantly, the way I think about economics. I am grateful for having the opportunity to work with them.

I would also like to thank my beloved parents, José Maria and Beatriz, who have always supported my choice to pursue graduate school abroad, even though that meant being away from them. Likewise, I would like to thank the support of my sister Paula and my brother Fernando. I carry their love with me wherever I go.

Coming abroad from Brazil to study at UCLA allowed me to meet a very special person. I would like to express my gratitude to Deb, whose continued emotional support pushed me forward when I thought it was not possible.

Many thanks to my friends Hadi, Jon, Tina, Keyoung, and Kun, who went through graduate school with me and became part of my extended family in Los Angeles.

This work would not have been possible without the help of Claudio Lucinda and João Machini. Claudio introduced me to the Brazilian data that was used in the first two chapters of this dissertation, and showed me each step of the process to access it. João Machini is a a dear friend, who is also an expert in computer science. He was kind and helpful, structuring the Brazilian data in a format that an economist could work with.

Finally, I would like to thank the Economics Department at UCLA for continued financial support as well as the amazing staff that supported me throughout the five years completing the program.

VITA

| | |
|---|---|
| 2006–2009 | B.S. (Mathematics), University of São Paulo, São Paulo, Brazil. |
| 2010–2012 | M.A. (Mathematical Economics), National Institute of Pure and Applied Mathematics (IMPA), Rio de Janeiro, Brazil. |
| 2013–2014 | Research Assistant, Economics Department, University of São Paulo, São Paulo, Brazil. |
| 2015–2019 | Teaching Assistant, Economics Department, UCLA. During summer 2017, taught Microeconomic Theory (Econ 11) as part of the Economics Summer Program. |
| 03.2017-08.2017 | Research Assistant, Anderson School of Management, UCLA. Project: "Acquiring Information Through Peers". Professors responsible for the project: Bernard Herskovic and João Ramos. |
| 01.2017-10.2017 | Research Assistant, Anderson School of Management, UCLA. Project "Taxation and Market Power in the Legal Marijuana Industry". Professor responsible for the project: Brett Hollenbeck. |
| 2018–2019 | Teaching Assistant Coordinator, Economics Department, UCLA. |

# CHAPTER 1

# Menu Choice and Market Structure: Selection in the Brazilian On-site Procurement Sector

## 1.1 Introduction

The trade-off between efficiency and equity is a central topic in economics. While the concept of efficiency is related to the optimal production and allocation of resources in the economy, equity is related to the distribution of these resources throughout society. Public contracting can be seen through the lens of both concepts. Unambiguously, the government cares about paying the smallest price for a given contract. At the same time, redistribution policies can increase the opportunity to win contracts to a broader set of firms, thereby promoting development across regions.

In the context of procurement auctions, efficiency means awarding contracts to the lowest-cost firms, while equity means spreading rents across a distinct set of firms. While the desired rent distribution a social planner might have is unclear, common policies such as bid-subsidies[1] and set-asides[2] to small firms demonstrate through revealed preference that policy makers care about equity.[3] Sometimes this is stated as a policy goal:[4]

---

[1]Bid subsidy provides discounts to qualified firms in terms of their bid. Usually, it is used for determining the winner while the contract itself is awarded at the actual bid amount. For reference, Krasnokutskaya and Seim (11) studied the design of bid subsidy programs in highway procurement auctions.

[2]Set-asides is the practice of limiting competition for certain contracts to a specific group of bidders, for example to small firms.

[3]Two examples of such policies are: the Small Business Act in the US and law 123/2006 that established differential treatment for micro and small firms in Brazil.

[4]Opening remarks in "Cartilha do Comprator: os novos paradigmas da administração pública", SEBRAE. The "new concept" refers to the differentiated treatment for Micro and Small Enterprises in Brazilian public

*"The new concept of government contracting allows the Federal Public Administration to encourage sustainable regional development, increase employment, increase competitiveness and bring improvements in income distribution. Public procurement is not just a process: it is an economic policy where the Federal Government uses its buying power to promote diverse business sectors."*[5]

*Loreni Foresti*

*Secretary of Logistics and Information Technology*

*13.03.2015, Ministry of Planning, Brazil*

In this chapter, I identify a new variable that influences the distribution of rents across firms: the composition (or menu) of items in a given auction session. Using a large data-set on Brazilian procurement auctions, I show evidence that entry is costly and that the mix of products being auctioned off is a first order effect to understand firm participation. In the following chapter, I continue the analysis with the focus on estimating entry costs and quantifying the degree of inefficiency present in this market.

Several products procured within the same location give participants an opportunity to bid for multiple items. The government will often host a session for a single item when large contracts are in play, such as in the construction sector and sales of mineral rights. By contrast, for smaller contracts - such as in the case of off-the-shelf goods and short term services, it is not rare to observe goods being bundled into the same session.[6]

From the auctioneer's perspective, there can be gains of scale associated with bundling items in a single session. This is the case when it is costly to host auctions, e.g. when an auxiliary facility has to be rented or additional employees needs to be hired. Thus, it might not be in the best interest of the auctioneer to offer a small contract of ink cartridges when it can a purchase a lot of office supplies simultaneously.

---

procurement.

[5]ibid. (My translation).

[6]Examples of multiple items sold at the same session are common. The reader might be familiar with auctions for art, stamps, cattle, timber, treasury bills and seized assets auctions.

The bundle also indirectly selects firms that can profit from entering a session. This matters when sellers are multi-product. In this case, they can benefit from the complementary goods being procured. For example, a firm that sells office supplies is more likely to participate in a session if the government is looking to buy both paper and ink cartridges rather than just paper. There is a trade-off between bearing fixed costs for entering a session and the menu of goods being auctioned. While the menu of items might increase a firm's expected profit from entering the session, fixed costs might prevent it from doing so in the first place. Thus, the extent to which the menu of items affects auction outcomes is an empirical question.

To quantify the effects of session composition in auction outcomes, I collect an extensive data-set from the on-site Brazilian procurement sector between 2003 and 2006. Similarly to the U.S., this sector accounts for a large fraction of GDP.[7] For example, in 2014 it accounted for 8.2% of Brazil's GDP while the global share was 7.7%, according to the OECD.[8] Since the early 2000s, the open descending price auction, also know as the Dutch auction, has been broadly used for the acquisition of shelf goods and services.[9] The main distinction from other existing auction procedures in Brazil is that the open descending format does not have a cap on the estimated value of contracts being auctioned.

The case of on-site procurement auctions in Brazil is an ideal setting to study how the menu choice shapes market outcomes. First, all information regarding the tender procedures are publicly available at the Brazilian procurement platform ComprasNet.[10] This translates into detailed bid level data, and allows me to track firm participation for each item purchased by the government within a session. Second, on-site auctions require an employee or

[7] The share in the U.S. is close to 10% according to Krasnokutskaya and Seim (10).

[8] www.oecd.org/gov/lac-brazil.pdf

[9] Brazil uses six types of distinct procedures to establish contracts with the private sector. The procedures differ in terms of their complexities and value of contracts involved. The procedures are: reverse auctions for purchases ("pregão"), invitation-based auctions ("carta-convite"), two types of sealed-bid auctions ("concorrência" and "tomada de preço"), auctions for sales of seized goods ("leilão") and contests ("concurso"). https://www.comprasgovernamentais.gov.br/images/conteudo/ArquivosCGNOR/SEBRAE/Compras-Pblicas.pdf, pp. 9-14.

[10] www.comprasnet.gov.br.

representative of a firm to be present at the auction location in order to submit a valid bid. Thus, entry costs are non-negligible and reinforces the effect that the menu has on entry. Third, the format of each auction follows a two-stage format with a first round of proposals and a second stage of open descending price auctions. Thus, the main action is at the entry decision. Bid functions do not have to be estimated in order to recover model fundamentals and it facilitates the introduction of asymmetry across bidders in a structural framework.

I find evidence that the data is consistent with a theory of selection. The average entrant has a higher portfolio match with the session and they are closer to auction locations. Distance affects entry decisions negatively: a 1 unit (100km) increase in the distance to the auction location lowers the odds ratio for entry 0.91 times. At the same time, an additional auction in the set of potential auctions of a firm increases the odds for entry 1.62 times.

In terms of variable costs, a 1% increase in the distance to auction location increases bids by 0.4% to 3.3%. There are also gains of scale in terms of the size of the contracts: a 1% increase in the contract quantity for a given product increases bids by 0.64% to 0.76%. While smaller firms seem to have a lower average cost than larger firms, this does not hold in terms of winning bids. Instead, the main force responsible for lowering procurement costs is the presence of additional bidders. I find that an extra bidder can lower costs between 21.2% and 32.4%. These results motivate and feed into the structural model presented in chapter 2 which I use to estimate firms' entry costs into multi-object sessions.

The rest of the chapter is structured as follows. The literature is reviewed in section 1.2. Section 1.3 describes the data and the letting process. Section 1.4 presents the stylized facts and the reduced-form analysis. Section 1.5 concludes.

## 1.2 Literature review

The present and the following chapter fill in a gap in the empirical industrial organization literature on multi-object auctions. To the best of my knowledge, this is the first attempt to empirically analyze multi-object auctions. Parallel to this setting, a few researchers have considered structural analysis of multi-unit auctions.[11] This is the case of Brendstrup & Paarsch (05) who study bidder asymmetries in multi-unit fish auctions in Denmark, and Hortaçsu & McAdams (10) who analyzes inefficiencies in Turkish Treasury auctions.

The big picture of this project relates to Bulow & Klemperer (09) and Roberts & Sweeting (13). They study a sequential mechanism, as an alternative to simultaneous auction, that can lead to an increase in surplus, allowing both buyers and sellers to gain. In contrast, I study the composition of sessions as a policy variable and show that it is also a channel for the redistribution of surplus *across* firms.

The framework that will be taken to the data departs from the literature on endogenous entry in auctions. I follow Gentry and Li (14) and Ye (07) who provide frameworks to study entry and identification of model primitives in single item auctions where entry is costly. Their work uses an affiliated signal approach (AS), where bidders have independent private values but only observe a noisy signal of their valuation prior to entry. The (AS) framework fits between the pivotal works of Samulelson (94, S) and Levin and Smith (94, LS). On one hand, the S model assumes the bidders know their exact valuation prior to entry. Thus, entrants should make positive profits on average. On the other hand, the LS model assumes entry decisions are made prior to learning any information on valuations. The implication that follows is that entry is random. I study on-site auctions where entry is inherently costly. Data patterns from Brazilian on-site auctions show that the selection process is not random but many bidders never win a single auction, suggesting that the (AS) model is suitable.

From the equilibrium of the model I derive moment inequalities on entry decisions that induce bound entry costs. I build on Tamer (2003), Ciliberto and Tamer (2009), among others using the moment inequalities approach for bounding parameters in games with entry. This method is convenient in my setting bidders' signals are not observed. Rather, I observe entry decisions. A distinction from my setting to the above mentioned work is that cost distributions are not point identified because the Brazilian auction format used is the open descending price and there are jump bids. I follow Haile & Tamer (03) who demonstrate how

[11] These two types of auction differ since in multi-unit auctions the aggregate number of units won matters, while in multi-object auctions each object matters.

to construct informative bounds on cost distributions using bid history and dropout points. Since the zero profit condition for marginal entrants relates truncated cost distributions with entry costs, I can construct moments that induce bounds on entry costs.

This project is one in a few that uses Brazilian procurement data on policy analysis. In his dissertation, Szerman (12) pioneered the use of Brazilian procurement auctions by studying set-asides programs for small firms, online auctions, and the transition from on-site to online auctions. Ferraz, Finan, and Szerman (16) analyzed the effects of procurement contracts on firm dynamics. Here, I use on-site auctions to study how the mix of products within a session and barriers of entry shapes market outcomes such as firm participation and procurement costs.

## 1.3 Data description

The on-site data was collected from the Brazilian federal purchasing platform ComprasNet. ComprasNet operates procurement processes and provides information regarding the contracts run by the Brazilian Federal Public Administration. On the platform, minutes, manuals, relevant legislation, and suppliers' characteristics can be consulted. In Brazil, the reverse price auctions were regulated in 2002 by law 10.520.[12] I focus on the window between 2003 and June 2005 as it is the earliest period where the data is available, on-site was the most common procedure to procure contracts, and there were no policies benefiting small firms.[13]

The minutes displayed detail the information on all items and services being purchased at each session, bidding data with timestamp and firm identification. The platform provides information on firm location, size, and line of products. Particularly, the measure of size is not ideal as it started to be displayed in the data in 2006, following the developments of laws that benefited small firms. To mitigate this issue, I only consider two categories for

[12] http://www.planalto.gov.br/ccivil_03/leis/2002/L10520.htm.

[13] On June 2005, the government regulated online auctions making the electronic format the preferred procedure. Law 5.450-05, Art. 4.

firm size: *small* if the firm is micro or small according to the government classification and *big*, if otherwise.[14]

### 1.3.1 Letting process: on-site sections

The letting process consists of two different phases: internal and external. The internal phase starts with purchasing units elaborating a document ("termo de referência") with the specification of objects and services to be contracted by the government as well as the team that will conduct the auctions. Once it is approved the process moves to the external phase.

Tender procedures are published at least eight business days prior to each public session. This information is always posted in the official government newspaper ("Diário Oficial da União") as well as on the internet.[15] For large contracts the information is also posted in newspapers. The tender procedures contain a summary of the description of the items, as well as the date, time, and location of the public session. Detailed information regarding each auctions is further available in a public notice ("edital").

Following the schedule, interested bidders attend the session, go through an accreditation process and submit their sealed proposals. The auctioneer ranks all proposals according to their values. Bidders that submitted prices up to 10% above the top ranked proposal move to the open descending price auction. If there are no three bidders that meet these conditions, the top ranked proposals (maximum of three) move to the next phase.

The open descending auction follows a sequential process starting from the firm with the highest proposal submitted. This corresponds to the worst proposal that moved to the open bid stage. It is asked if it is willing to beat the lowest outstanding bid. If the answer is negative, it drops out of the auction and cannot bid any further. If the answer is positive, they place a bid and the outstanding price is updated. The auctioneer moves to the second

[14] In 2006 the government defined labels for firms according to their gross revenue: micro if below R$360.000; small if between R$ 360.000,01 R$ 3.600.000,00 and other (medium/large) if above R$ 3.600.000,01. For reference, the nominal exchange rate in 01.04.2010 was 1.72 Brazilian Real to One U.S. Dollar. http://www.planalto.gov.br/ccivil_03/leis/LCP/Lcp123.htm.

[15] Law $n^o$ 10520/2002.

highest firm ranked and so on. This process goes on until all but one firm drops out, who is awarded the contract with a value corresponding to the last outstanding bid.[16] After all auctions end, the auctioneer verifies and homologates the winner of each item, canceling a contract if there are irregularities. Contracts are signed and the minutes containing all tender procedures are made available online. A snapshot of the data is provided in figures 1.1 and 1.2.

### 1.3.2 Descriptive Statistics

In this section I describe the variables that were collected, starting from the macro variables in terms of total number of auctions and market shares to the micro level, analyzing bid level data and product characteristics. They will serve as an input for a reduced-form analysis and will help motivate the structural model. Table 1.1 lists the variables observed in the data.

A summary, displayed in table 1.2, shows that the collected data contains approximately thirteen thousand sessions that took place between 2003 and 2006 and eighty nine thousand valid auctions for off-the-shelf goods and services. They consist of purchases at the federal level with a few examples of buyers being federal hospitals, federal universities, and the army. Auctions within a session are either exclusively composed by off-the-shelf products (61%) or exclusively by services (38%). I concentrate the analysis on off-the-shelf sessions with total value between $1,000 and $10,000,000 and auctions with value above $100, which restricts the data to approximately fifty thousand auctions and eight thousand sessions. Unless stated otherwise, auctions for services will not be used.

The time series presented below in figure 1.3 suggests that there is seasonality in the total number of sessions held through time, with peaks at the end of each calendar year. This is also verified in the time series of online auctions (figure 1.4). While it does not come as a surprise, since the federal budget is renewed on an annual basis, it contrasts with two relevant economic concepts. First, there is a value on having flexibility purchasing goods or

[16] There are no regulations on bid jumps, i.e., by how much a firm has to undercut each other proposal.

hiring a service without too much delay. If this was the main force in play, one should expect a more smooth time series. Second, if planning and inventory were the main considerations, the clustering of sessions would likely occur at the beginning of the calendar year and not at the end.

There are approximately seven hundred federal units that buy from ten thousand firms. Buyers are concentrated in the southeast region and state capitals, places responsible for the greatest share of Brazilian GDP (figures 1.5 and 1.6).[17] The top five buyers are all from Rio de Janeiro (former federal district) and Brasilia (current federal district), locations known for a strong federal presence as shown in table 1.4. Sellers follow a similar pattern (figure 1.7).

A good feature of the data is that the product space is unusually well defined. This is helpful because there is no need to work with strings or run a product classification algorithm besides being convenient to control for product fixed effects. Two variables are useful for this purpose. One is a six-digit product specific code. The second is a product class number that consists of four digits, the first two referring to a broader group the product belongs to. Overall, I observe 438 product classes and 77 product groups. The most common classes of goods purchased are equipment for data processing, equipment (17.3%), followed by manufacturing goods 8.3% components for electronic equipment (6.9%) and office supplies (6.7%) as depicted in table 1.3.

Sessions are indeed multi-object and firms do participate in multiple auctions. According to table 1.5, in an average off-the-shelf session approximately 8.9 items are procured to 4.3 firms that participate in 5.4 items each, each auction with 2.34 participants. The services sector is quite different with an average of 2 auctions procured in each session. This table is also informative that the level of competition in each individual auction is not high. This suggests that the auction framework is still appropriate for modeling purposes, contrasting the case of auctions with many participants that can be thought as a competitive environment.

[17]https://www.ibge.gov.br.

## 1.4 Stylized Facts

In this section I address the main facts from the data. First, I discuss the two-step design and show how it differs from standard open descending price auctions. Second, evidence of selection both in terms of entry cost and items procured is presented. This is a key ingredient that motivates the structural model. Third, market outcomes such as firm size distribution and market shares are analyzed. Lastly, I quantify the extent to which session covariates and firm characteristics induce entry.

### 1.4.1 Two-step design

As discussed in subsection 1.3.1 three bidders are always guaranteed in the open descending phase of the auction. Therefore, the first round only bites when there are more than three players present. In this case, the first round selects some of the bidders and sets a reserve price for the following stage.

Szerman (12) showed that when players know their exact valuation at the time of submitting a proposal there is no symmetric, strictly increasing equilibrium bidding function in the first stage. An assumption here is that players would still go through the competition process on the second round. Complementing his analysis, there are other equilibria in this setting. In a strictly increasing equilibrium bidding function, players that move to the second round learn their opponents' value. Thus, lower ranked players update that they will get a zero payoff and become indifferent between any drop out point. The limit case is when they drop out immediately, which makes the outcome equivalent to a first price auction. Other types of equilibria involve weakly monotone strategies, such as in the case where all players bid at the upper boundary of the support. This is an equilibria because the best three players always move to the second round so it does not pay off for a single bidder to deviate. In any case, any equilibrium bidding function in this setting is weakly monotone on values because single crossing still holds. This is enough for the empirical analysis, since the identification strategy of value distributions relies on order statistics.

Empirically, many reasons might induce rank inversions between the first and second stage of the auction. Some of them are: firm inattention (not perceiving competition), overconfidence (believe value is higher than it truly is), learning (proposal based on a signal), and winner's curse (common value component realized in the end of first round). In the data inversions do happen. Ranking inversions have a higher frequency in auctions with more than three bidders (table 1.6, 41.3%) compared to auctions with a maximum of three bidders (table 1.7, 26.4%). The difference can be attributed in part to the fact that 1.6 contain the cases where bids were at least 10% close to the best proposal submitted.

In order to analyze the role of the two-step design on auction outcomes, I ran a robustness check on the structural estimates from chapter 2. I used auctions with up to three bidders and auctions with up to ten bidders for the estimation and there was no distinction on the results. Throughout the rest of chapters 1 and 2, I hold the assumption that the rankings from the second stage reflect the true order of valuations in each auction.

### 1.4.2 Evidence of selection

On-site sessions requires the physical presence of bidders to the session location as well as researching costs for each item they would like to bid for. This, per se, suggests that entry costs are present. A theory of selection based on entry costs has the empirical prediction that firms would stay out if entry costs are high enough and, conditional on entering, they should participate in as many items as possible - given that each auction is costless. On the other hand, no selection on entry costs would imply that participation in any item is profitable and variables such as distance or contract size should not play a role in entry decisions.

An important driver of entry is the distance to the session location. Figure 1.8 illustrates that it has heterogeneous effect on small and big firms where the latter is willing to bid in locations that are further away while smaller firms bid in closer auctions. This variable is likely to be correlated with both entry costs (mandatory on-site presence) and variable costs (transportation cost/delivery). While entry costs are unobserved the latter point is verified in the reduced form analysis of Section 1.5.

Another relevant fact is that firms are multi-product. On average they have 17.7 product classes (standard deviation of 47.2), ranging from 1 to 310 classes. Having their portfolio of products in hand, I go through each session and compute the number of their potential auctions for each session (auctions they could have participated in) and contrast it with the actual number they bid for. Following Roberts and Sweeting (13), I define the set of potential bidders for a session as the collection of firms that 1) actually participated in the corresponding session, 2) have at least one item of interest and 3) either participated in the respective location in the past or participated in another location 300km close by.[18]

I find that conditional on entering the session, firms participate in a large share of their potential auctions. This is shown in figure 1.9. Moreover, in terms of entry decisions, firms that chose to enter had a higher match with the products being purchased, at the session: 69% match for entrants versus 32% for non entrants. This result also holds in terms of share of potential revenue (81% match for entrants versus 39% for non entrants) and number of auctions (11.15 for entrants versus 2.73 for non entrants). Lastly, the average entrant is halfway closer to the session relative to the non entrant firm: 344km versus 625km, respectively. These findings are presented in table 1.8.

### 1.4.3 Firm size distribution

Contracts are unequal in size. This is specially true for the case of off-the-shelf goods (figure 1.10). Small firms were awarded (51.3%) of the 18,554 off-the-shelf contracts with value between \$1,000 and \$10,000.[19] These rents are not homogeneously distributed within firms. The Gini coefficient of rent distribution is 0.6 with similar results breaking it down by firm size and value of the contract. The precise distribution of procurement costs can be seen in figure 1.11.

The multi-object nature of the sessions plays an important role. Conditional on winning,

[18] Note that this definition does not impose a hard constraint on how close by a firm has to be because it is in terms of *buyer location* and not *seller location*.

[19] There are 2,768 small firms and 542 big firms. I restrict the sample to make a fair comparison of the rents obtained, taking into account that contracts are indivisible and heterogeneous in value.

firms are awarded multiple contracts in the same session, ranging from 1 to 118 contracts and an average of 3.4 contracts per session.

Even though contracts seem evenly distributed among small and big firms, this is not the case when the size of the contract is considered. In the sequence of figures 1.12-1.23 a series of graphs related to market shares are displayed. They are computed separately in terms of the number of contracts (equally weighted) and the value of the contracts awarded (value-weighted). I plot these graphs for off-the-shelf goods and services according to firm size, distance, and firm size/distance. The key takeaway is that the market shares are distorted in favour of big and non-local firms when the value of the contract is taken into account. That is, different types of firms win contracts in similar rates, however, big firms and firms further away win the largest ones.

## 1.5 Drivers of entry and procurement cost

To quantify the extent to which firm size, distance, and the portfolio match between a firm and the session influences entry decisions at the session level, I take advantage of the fact that firm locations and their line of products are observed. Focusing on the set of potential firms for each auction session, I specify a fixed effects logit model of entry decisions based on session characteristics.[20] That is, a firm $i$ in session $t$ will participate as long as it is profitable to do so.

$$Entry_{it} = \begin{cases} 1 \text{ if } \alpha_i + x'_{it} \cdot \beta + u_i > 0, \\ 0, otherwise \end{cases}$$

where $u \sim F(\cdot)$, $F(z) = \frac{exp(z)}{1+exp(z)}$ and $x_{it}$ includes firm size, distance, and measures of portfolio match between firm $i$ and session $t$.

[20]The conditional likelihood is used in order to avoid the incidental parameter problem.

The results displayed in table 1.9 provide evidence that firms that are further away are less likely to participate while firms that have a higher portfolio match are more likely to do so. For convenience of interpreting the results I fix the number of potential auctions in a session as the measure of portfolio overlap. Ceteris paribus, a 1 unit (100km) increase in the distance to the session location lowers the odds ratio for entry by 0.91 times. At the same time, an additional auction in the set of potential auctions of a firm increases the odds for entry by 1.62 times. The results on firm size are not conclusive.

The empirical observation presented in table 1.8 shows that some firms could, but chose not to, participate in auction sessions, suggesting there is selection in the observed bid distribution. To control for selection and analyze the underlying cost distribution according to firm type, I specify a Heckman selection model of bids on auction characteristics, where bids are observed if it was profitable for firms to enter the corresponding session. That is,

$$Bid_i = \begin{cases} x_i' \cdot \beta + u_i \text{ if } z_i'\gamma + v_i > 0, \\ 0, otherwise \end{cases} \tag{1.1}$$

where

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} \Bigg| x_i \sim N\left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_u^2 & \rho\sigma_u\sigma_v \\ \rho\sigma_u\sigma_v & \sigma_v^2 \end{bmatrix} \right) \tag{1.2}$$

$x_i$: quantity supplied, firm size, and distance.

$z_i, x_i$: portfolio match with the session and interaction between firm size and distance.

The underlying assumption of this selection model is that, conditional on observables, the portfolio match is a variable that only shifts entry decisions but does not shift the bids submitted in the auction. This matches the intuition from economic theory. Once the cost of entry is sunk, bidders are expected to bid until the value of the contract reaches their own cost.

The findings are summarized in table 1.10. Larger contracts, a better product match with the session, and small firms, predicts a more likely participation, while distance to the auction location decreases the likelihood of participation, which has a stronger effect on smaller firms.[21]

In terms of the underlying cost distributions, larger contracts predict higher costs to the firms, but with gains of scale. A 1% increase in the contract quantity increases the cost of firms by 0.64% to 0.76%. Higher distances also increase firms' costs: a 1% increase in the distance to the auction location increases costs by 0.4% to 3.2%. Finally, the average cost for a small firm ranges from 5.7% to 32.3% less than for a big firm. Because the tail of the distribution matters more than the average bid for winners, this last result does not necessarily translate in winners that are small firms having lower costs than winners that are big firms. As seen below, this is actually a feature of the data.

In table 1.11, I estimate the drivers of procurement cost instrumenting for auction participation with the firm portfolio match. Controlling for a rich set of fixed effects including time, location, product and supply type, I find three important results. First, the participation is the main driver of lowering the costs of the contract: the presence of an additional bidder has the effect of bringing down contract costs by 21.2% to 32.4%, everything held constant. This is an important result that illustrates the power of competition when entry is selective and goes in line with Bulow & Klemperer and Robert & Sweeting. In a setting where entry is costly enough so that only a few bidders show up for the auction, keeping participation high is key to keep procurement costs low.

Second, there are gains of scale in terms of contract size. A 1% increase in the quantity supplied increases the cost of the contract per quantity by 0.88% to 0.90%. When bundling contracts is possible, it would not only lower costs per unit due to gains of scale but would also increase the size of the contract, thereby providing incentives for more firms to participate.

[21] I do not take a strong stance analyzing the selection part of the Heckmann model since fixed effects are only included in the second stage, due to the incidental parameter problem. This and the different assumptions on the distribution of the unobserved term, explains the differences between table 1.9 and the selection part of table 1.10

Third, distance impacts the firm's variable costs.[22] An extra 100km of distance to the auction location is predicted to increase contract costs by 1.28% to 3.03%. Thus, firms that are further away face both lower odds of entry and higher costs for the contract. I do not find a strong association between firm size and the cost of the contracts in play.

## 1.6 Conclusion

In this chapter I studied the case of the public on-site procurement sector in Brazil, which features multiple objects being sold in the same location. Focusing on the data between years of 2003 and 2006, I showed that firms do respond to the design of the section for entry in terms of contract size, location, and product selection.

A strong selection process is created as procurement costs fall substantially given extra bidder participation. To perform this analysis, I used the portfolio match between firm and section as an instrument for participation. That is, conditional on observables the match is a driver of entry but does not affect the underlying cost or bid behaviour of the firm. Such assumption rules out potential synergies between products that might decrease the firm's costs and it is a subject of future work.

These results motivate and feed into the structural model presented in the following chapter, where heterogeneous firms trade-off their expected revenue from entry against entry costs. Firms differ in terms of their line of products and location. This heterogeneity differentiates them in terms of profits and fixed costs allowing me to study the barriers of entry that exists in this market according to firm type.

[22] Most likely due to transportation costs and delivery.

## 1.7 Appendix I: Tables

| Type | Variable | Description |
|---|---|---|
| Buyer (government) | buyer id | 6 digit identifier |
| | location | latitude and longitude |
| Seller (firm) | firm id | 14 digit identifier |
| | location | latitude and longitude |
| | size | categorical: micro, small, other |
| | product line | products eligible for supply at *ComprasNET* |
| Session Characteristics | session id | 12 digit identifier. Also identifies buyer |
| | products listed | product code, class, and group |
| | date | |
| Auction Characteristics | session id | session it is part of |
| | item listed | product code, class, and group |
| | supply unit | |
| | quantity | |
| Bid Information | proposal | phase 1 proposals with firm id |
| | bid | phase 2 bids with firm id and timestamp |
| | winner id | winner firm id |
| | status | homologation of auction |
| Item characteristics | product code | 6 digit identifier |
| | product class | 4 digit identifier |
| | product group | 2 digit identifier |
| | type of product | dummy for off-the-shelf or service |

Table 1.1: List of variables.

| Statistic | Raw Data | Final Sample |
|---|---|---|
| Sessions | 12,596 | 8,065 |
| Buyers (government) | 978 | 726 |
| Sellers (firms) | 13,949 | 9,974 |
| Auctions | 93,940 | 54,678 |
| Valid auctions | 88,504 | - |
| Cancelled auctions | 5,436 | - |
| Item Classes | 474 | 438 |
| Item Groups | 77 | 77 |

Table 1.2: Statistics - all Sessions. Final sample consists of all sessions with value between $1,000 and $10,000,000 and individual auctions with value above $100.

| Name | Number of auctions |
|---|---|
| Equipment for data processing | 15,306 |
| Manufactured goods, non-metallic | 7,370 |
| Components for electrical/electronic equipment | 6,141 |
| Office supplies | 5,991 |
| Furniture | 2,903 |

Table 1.3: Top 5 product classes for all on-site sessions. A product class is defined by a four digit code where the first two numbers represent the product group it belongs to.

| Buyer | Number of sessions |
|---|---|
| District Court (DF) | 146 |
| Military Court (DF) | 147 |
| Health organization (RJ) | 178 |
| $2^o$ Batalhão Ferroviário (RJ) | 191 |
| Oswaldo Cruz Foundation (RJ) | 398 |

Table 1.4: Top 5 Buyers. A buyer is an administrative unit of general service (UASG).

| Contract Type | Level | Variable | Mean | Std | Min | Max | 5% | 95% |
|---|---|---|---|---|---|---|---|---|
| Off-the-shelf | Session | Total Value | 165,377.5 | 614,366 | 1003.09 | 9,956,734 | 3,597.89 | 606,058.8 |
| | | Entrants | 5.35 | 4.77 | 1 | 38 | 1 | 15 |
| | | Number of items procured | 9.33 | 15.02 | 1 | 329 | 1 | 34 |
| | | Number of auctions per firm | 6.37 | 8.74 | 1 | 130 | 1 | 22 |
| | | Observations | 5,251 | - | - | - | - | - |
| | Auction | Value | 17,105.25 | 119,945.8.8 | 100.02 | 8,486,696 | 136.28 | 56,351.8 |
| | | Entrants | 3.9 | 3.06 | 1 | 26 | 1 | 10 |
| | | Observations | 49,005 | - | - | - | - | - |
| Services | Session | Total Value | 185,303.8 | 586,950.1 | 1,027.77 | 8,490,74 | 2,314.01 | 800,794.1 |
| | | Entrants | 4.38 | 4.6 | 1 | 50 | 1 | 13 |
| | | Number of items procured | 2.02 | 3.28 | 1 | 52 | 1 | 6 |
| | | Number of auctions per firm | 1.74 | 2.42 | 1 | 46 | 1 | 5 |
| | | Observations | 2,814 | - | - | - | - | - |
| | Auction | Value | 91,694.56 | 3,372,659.4 | 103.44 | 8,240,209 | 373.64 | 365,893.4 |
| | | Entrants | 4 | 4.25 | 1 | 50 | 1 | 12 |
| | | Observations | 5,673 | - | - | - | - | - |

Table 1.5: Overview of auction outcomes. Values adjusted for inflation. Base is Jan/2010. Sample consists of all sessions with value between $1,000 and $1,000,000 and auctions with value above $100.

| | | 1 | 2 | 3 | 4 | 5 | 6+ | Total |
|---|---|---|---|---|---|---|---|---|
| | | $2^{nd}$ phase | | | | | | |
| $1^{st}$ phase | 1 | 7,375 | 1,675 | 920 | 151 | 44 | 10 | 10,175 |
| | 2 | 2,294 | 6,776 | 930 | 156 | 23 | 19 | 10,198 |
| | 3 | 2,295 | 1,785 | 6,020 | 163 | 32 | 15 | 10,310 |
| | 4 | 2,358 | 1,793 | 919 | 5,174 | 32 | 12 | 10,288 |
| | 5 | 1,402 | 1,058 | 560 | 114 | 3,154 | 11 | 6,299 |
| | 6+ | 2,362 | 1,794 | 928 | 245 | 112 | 5,783 | 11,224 |
| | Total | 18,086 | 14,881 | 10,277 | 6,003 | 3,397 | 5,850 | 58,494 |

Table 1.6: Rank inversion between first and second phases of the auctions. Sample consists of all 21,602 off-the-shelf auctions with more than three participants.

| | | 1 | 2 | 3 | Total |
|---|---|---|---|---|---|
| | | $2^{nd}$ phase | | | |
| $1^{st}$ phase | 1 | 16,091 | 2,491 | 492 | 19,074 |
| | 2 | 4,271 | 8,714 | 472 | 13,457 |
| | 3 | 1,495 | 1,008 | 3,705 | 6,208 |
| | Total | 21,857 | 12,213 | 4,669 | 38,739 |

Table 1.7: Rank inversion between first and second phases of the auctions. Sample consists of all 27,403 off-the-shelf auctions with less than or equal to three participants.

| | *Entry* | | | |
|---|---|---|---|---|
| | No | | Yes | |
| | Mean | Std. | Mean | Std. |
| Potential Auctions (%) | .32 | .38 | .69 | .35 |
| Potential Value (%) | .39 | .42 | .81 | .31 |
| Potential Auctions (#) | 2.73 | 5.97 | 11.15 | 17.85 |
| Distance | 6.25 | 5.81 | 3.44 | 5.77 |
| Observations | 3,090,042 | | 26,558 | |

Table 1.8: Outcomes from entrants compared to firms that did not participate. All of-f-the-shelf sessions in the sample. Distance in multiples of 100km.

## 1.8 Appendix II: Reduced-form results

| | *Dependent variable: Entry* $\in \{0,1\}$ | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Potential Auctions (%) | 5.31*** | | |
| | (0.115) | | |
| Potential Revenue (%) | | 4.49*** | |
| | | (0.108) | |
| Potential Auctions (#) | | | 0.14*** |
| | | | (0.011) |
| Small | -0.046* | -0.044* | -0.033 |
| | (0.018) | (0.018) | (0.018) |
| Distance | -0.095*** | -0.096*** | -0.095*** |
| | (0.004) | (0.004) | (0.004) |
| Log Likelihood | -101,335.1 | -101,450.7 | -107,801.4 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.9: Conditional fixed effects logistic regression of entry decisions on session characteristics. Sample consists of potential participants for off-the-shelf sessions. Total number of observations in the sample is 3,116,600. Small is a dummy variable with value equal to one if the firm is *small*.

| | (1) Log Bid | (2) Log Bid | (3) Log Bid | (4) Log Bid | (5) Log Bid | (6) Log Bid |
|---|---|---|---|---|---|---|
| **Log Bid** | | | | | | |
| Log Quantity | 0.76*** | 0.67*** | 0.68*** | 0.76*** | 0.64*** | 0.67*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Small | -0.063*** | -0.29*** | -0.27*** | -0.059*** | -0.39*** | -0.28*** |
| | (0.013) | (0.013) | (0.013) | (0.013) | (0.014) | (0.014) |
| Distance | 0.0074*** | 0.024*** | 0.022*** | 0.0044*** | 0.032*** | 0.023*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Constant | 6.64*** | 8.46*** | 8.23*** | 6.38*** | 9.49*** | 8.40*** |
| | (0.022) | (0.021) | (0.022) | (0.024) | (0.037) | (0.038) |
| **Select** | | | | | | |
| Log Quantity | 0.019*** | 0.082*** | 0.090*** | 0.013*** | 0.074*** | 0.077*** |
| | (0.001) | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) |
| Small | 0.19*** | 0.15*** | 0.17*** | 0.19*** | 0.21*** | 0.21*** |
| | (0.004) | (0.005) | (0.005) | (0.004) | (0.003) | (0.004) |
| Distance | 0.0002 | -0.0028*** | -0.00068 | -0.0049*** | -0.0099*** | -0.0093*** |
| | (0.0004) | (0.000) | (0.001) | (0.000) | (0.000) | (0.000) |
| Small#Distance | -0.011*** | -0.010*** | -0.015*** | -0.0087*** | -0.0088*** | -0.010*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Pot. Auctions (#, 6 digits) | 0.077*** | | | | | |
| | (0.000) | | | | | |
| Pot. Auctions (%, 6 digits) | | 3.20*** | | | | |
| | | (0.005) | | | | |
| Pot. Revenue (%, 6 digits) | | | 3.21*** | | | |
| | | | (0.005) | | | |
| Pot. Auctions (#, 4 digits) | | | | 0.046*** | | |
| | | | | (0.000) | | |
| Pot. Auctions (%, 4 digits) | | | | | 1.56*** | |
| | | | | | (0.004) | |
| Pot. Revenue (%, 4 digits) | | | | | | 1.83*** |
| | | | | | | (0.005) |
| Constant | -1.92*** | -2.77*** | -3.09*** | -1.84*** | -2.60*** | -2.95*** |
| | (0.004) | (0.006) | (0.007) | (0.004) | (0.005) | (0.006) |
| $\rho$ | 0.40 | -0.11 | -0.021 | 0.38 | -0.26 | -0.046 |
| | (0.003) | (0.004) | (0.005) | (0.002) | (0.006) | (0.007) |
| $\sigma$ | 2.80 | 2.79 | 2.78 | 2.83 | 2.84 | 2.78 |
| | (0.008) | (0.011) | (0.013) | (0.009) | (0.018) | (0.019) |
| Observations | 1632952 | 1632952 | 1632952 | 1632952 | 1632952 | 1632952 |
| Censored Observations | 1431617 | 1431617 | 1431617 | 1431617 | 1431617 | 1431617 |
| Uncensored Observations | 201335 | 201335 | 201335 | 201335 | 201335 | 201335 |

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.10: Heckman selection model.

Table 1.10 refers to a Heckman selection model. Sample consists of potential bidders (entrants and non-entrants) for off-the-shelf auctions. Distance in per 100km units. Small is a dummy variable with value equal to one if the firm is *small*. Bids are in total dollar value submitted for the contract and adjusted for inflation (base is Jan/2010). Data restricted to the smallest bids submitted by each player in a given auction. Fixed effects for product, supply type, and auction are included on the first stage regression.

| | (1) Log Bid | (2) Log Bid | (3) Log Bid |
|---|---|---|---|
| **Log Bid** | | | |
| Participation | -0.391*** | -0.238*** | -0.234*** |
| | (0.091) | (0.087) | (0.087) |
| Log Quantity | 0.898*** | 0.875*** | 0.877*** |
| | (0.030) | (0.027) | (0.044) |
| Distance | 0.0286*** | 0.0128* | 0.0303*** |
| | (0.006) | (0.007) | (0.012) |
| Small | | -0.433*** | -0.273 |
| | | (0.109) | (0.268) |
| Small#Log Quantity | | | -0.00360 |
| | | | (0.033) |
| Small#Distance | | | -0.0265** |
| | | | (0.013) |
| Constant | 3.107*** | 3.723*** | 3.621*** |
| | (0.891) | (1.029) | (1.036) |
| Observations | 50,934 | 33,153 | 33,153 |
| R-squared | 0.595 | 0.638 | 0.639 |

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.11: Two-stage least squares regression of winning bids on observables. Distance in 100km units. The variable “Log Bid” corresponds to the procurement cost of the contract. Values in $ adjusted for inflation. Base is Jan/2010. Specification includes product, supply type, and location fixed effects as well as clusters for location and product. The variable ‘Participation’ is instrumented with the variable ‘portfolio match‘ which uses 6 digit product codes. Small is a dummy variable with value equal to one if the firm is *small*.

## 1.9 Appendix III: Figures

**Item nº** 2

**Descrição:** TOALHA

**Descrição Complementar:** TOALHA, MATERIAL PAPEL ALTA ALVURA, TIPO FOLHA 2 DOBRAS, COMPRIMENTO 23 CM, LARGURA 23 CM, COR BRANCA

**Quantidade:** 1000 **Unidade de Fornecimento:** PACOTE 1.250,00 FL

**Situação:** Aceito e Habilitado

**Aceito para:** UNIDAS COMERCIAL E DISTRIBUIDORA LTDA, **por R$** 9.680,00

Figure 1.1: Item 2 description (paper towel) of session 1/2005 held by the National School of Public Administration in Brasilia, buyer code 114702. Session was held at 03/14/2005.

**Item 2**

**Observação :** Participaram deste item os licitantes abaixo selecionados, com suas respectivas propostas. As propostas assinaladas com (*) foram classifcadas e os proponentes convocados para a fase de lances. As propostas assinaladas com (D) foram desclassificadas pelo Pregoeiro.

**Propostas**

| CNPJ/CPF | Razão Social/Nome | Qtde | Valor (R$) | |
|---|---|---|---|---|
| 04927672000106 | SCM COMERCIO LTDA | 1000 | 24.500,00 | |
| **Descrição Complementar:** OK | | | | |
| 37369576000130 | AMG DESCARTAVEIS LTDA | 1000 | 14.940,00 | |
| **Descrição Complementar:** OK | | | | |
| 04720671000188 | BAND DISTRIBUIDORA DE EMBALAGENS E DESCARTAVEIS LTDA | 1000 | 14.270,00 | |
| **Descrição Complementar:** OK | | | | |
| 02723637000131 | GUERRA PRODUTO DE LIMPEZA LTDA | 1000 | 13.120,00 | |
| **Descrição Complementar:** OK | | | | |
| 06910908000119 | jb brilante comercial ltda | 1000 | 12.350,00 | |
| **Descrição Complementar:** OK | | | | |
| 03449625000123 | MAXCLEAN COMERCIO E SERVIÇO LTDA | 1000 | 12.050,00 | |
| **Descrição Complementar:** OK | | | | |
| 04495604000107 | ELA DISTRIBUIDORA LTDA | 1000 | 11.940,00 | |
| **Descrição Complementar:** OK | | | | |
| 02982834000175 | MULTIPAPEIS DISTRIBUIDORA DE PAPEIS LTDA | 1000 | 11.000,00 | * |
| **Descrição Complementar:** OK | | | | |
| 03383327000188 | UNIDAS COMERCIAL E DISTRIBUIDORA LTDA | 1000 | 10.890,00 | * |
| **Descrição Complementar:** OK | | | | |
| 05972416000195 | NEW CLEAR COMERCIO DE PRODUTOS DE LIMPEZA LTDA | 1000 | 10.200,00 | * |
| **Descrição Complementar:** OK | | | | |
| 64088214000144 | TERRÃO COMERCIO E REPRESENTAÇÃO LTDA | 1000 | 10.140,00 | D |
| **Descrição Complementar:** OK | | | | |
| | aMOSTRA NÃO ATENDE O EDITAL | | | |

**Lances**

| Valor (R$) | CNPJ/CPF | Data/Hora do Lance |
|---|---|---|
| 10.150,00 | 02982834000175 | 14/03/2005 17:53:58 |
| 10.149,00 | 03383327000188 | 14/03/2005 17:56:32 |
| 9.800,00 | 05972416000195 | 14/03/2005 17:56:45 |
| Desistiu | 02982834000175 | 14/03/2005 17:56:52 |
| 9.790,00 | 03383327000188 | 14/03/2005 17:57:09 |
| 9.689,00 | 05972416000195 | 14/03/2005 17:57:24 |
| 9.680,00 | 03383327000188 | 14/03/2005 17:57:39 |
| Desistiu | 05972416000195 | 14/03/2005 17:57:45 |

**Aceitação**

| Data/Hora | Ocorrências | CNPJ/CPF |
|---|---|---|
| 14/03/2005 18:03:08 | Aceito | 03383327000188 |

**Habilitação**

| Data/Hora | Ocorrências | CNPJ/CPF |
|---|---|---|
| 14/03/2005 18:03:35 | Habilitado | 03383327000188 |

Figure 1.2: Item 2 description (paper towel) of session 1/2005 held by the National School of Public Administration in Brazilia, buyer code 114702, on 03/14/2005. The bid history is separated by the proposals phase “Propostas” and bid phase “Lances”. The variable “CNPJ/CPF” identifies a firm. Firms with “D” next to their proposal were disqualified. Firms with “*” next to their proposal moved to the open descending phase which is identified by “Lances”. Dropout points are market with “Desistiu”.

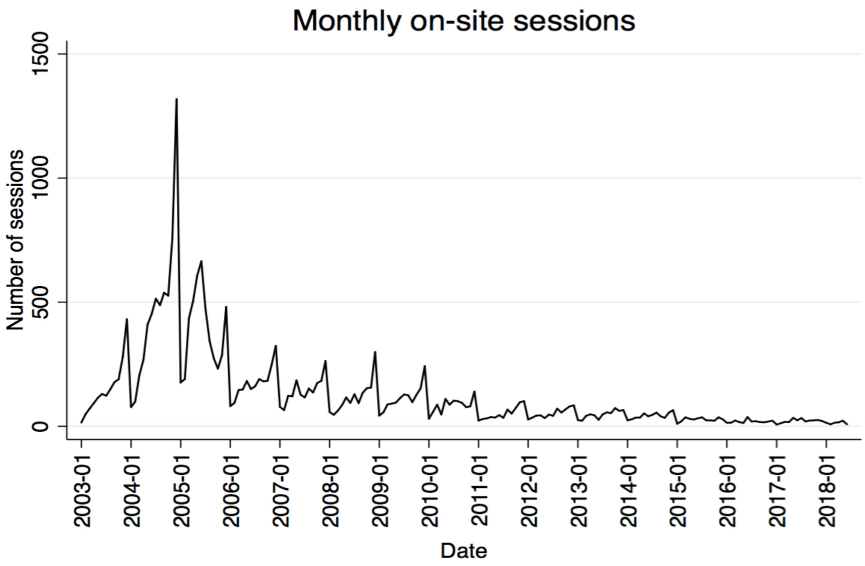



Figure 1.3: Monthly number of sessions. Sample consists of all on-site sessions between 2003 and 2018 regarding federal purchases of services and off-the-shelf goods, procured through open descending price format ("pregao"). On June 2005, the government regulated online auctions making the electronic format the preferred procedure for the federal acquisition of this type of contract. Law 5.450-05, Art. 4.

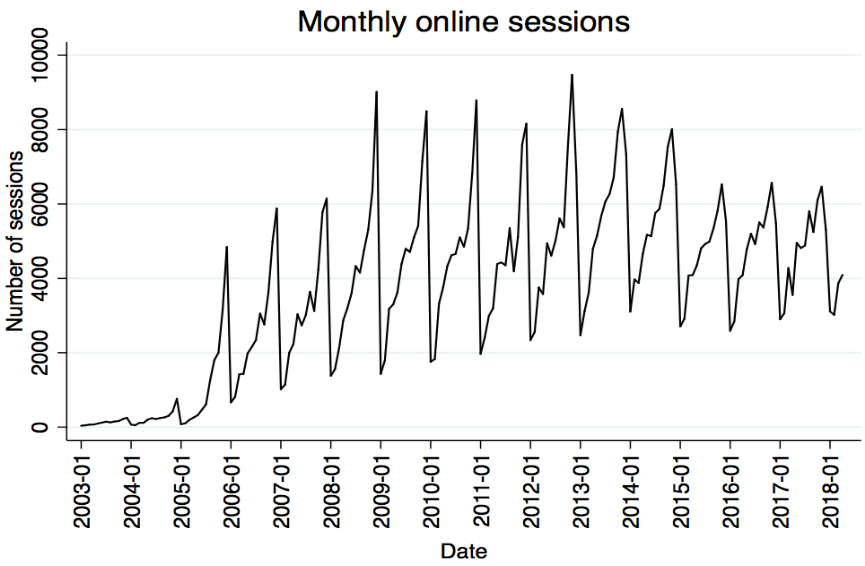



Figure 1.4: Monthly number of sessions. Sample consists of all online sessions between 2003 and 2018 regarding federal purchases of services and off-the-shelf goods, procured through open descending price format ("pregao"). On June 2005, the government regulated online auctions making the electronic format the preferred procedure for the federal acquisition of this type of contract. Law 5.450-05, Art. 4.

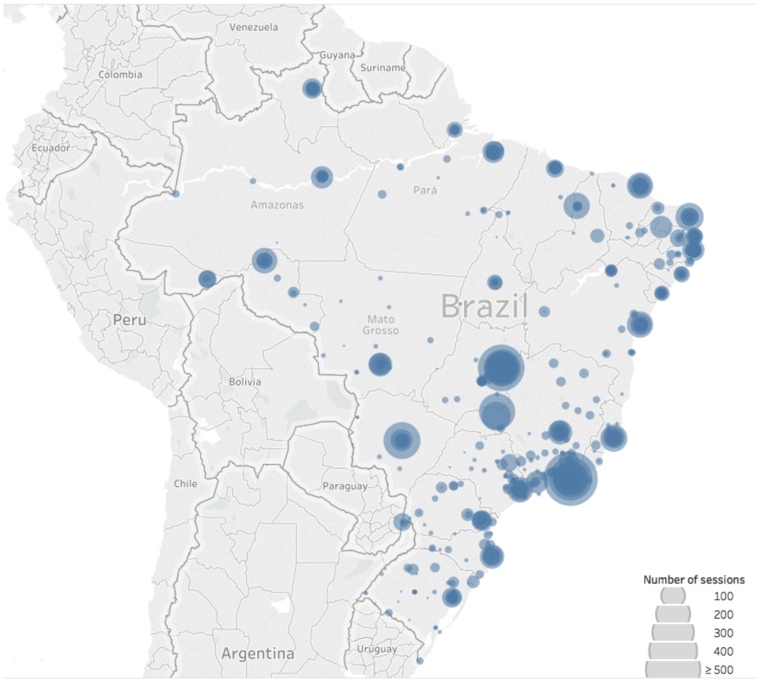



Figure 1.5: Total number of sessions by buyer location between 2003 and 2006. A buyer is an administrative unit of general service (UASG).

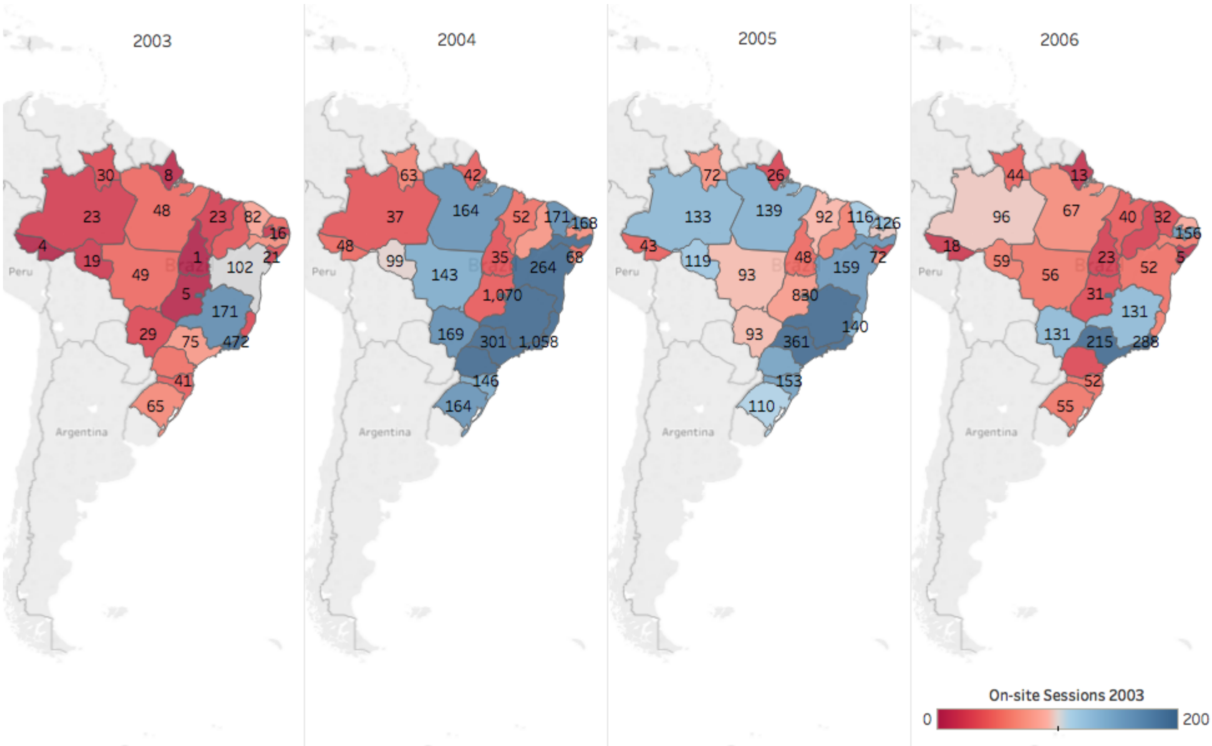



Figure 1.6: Total number of sessions by state in the years of 2003, 2004, 2005, and 2006.

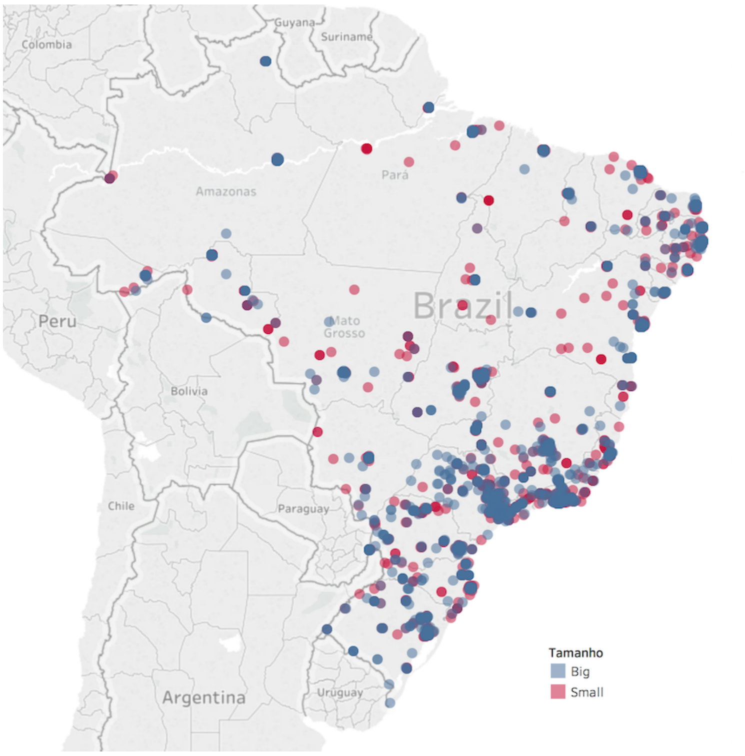



Figure 1.7: Seller location. Small firms in red and big firms in blue.

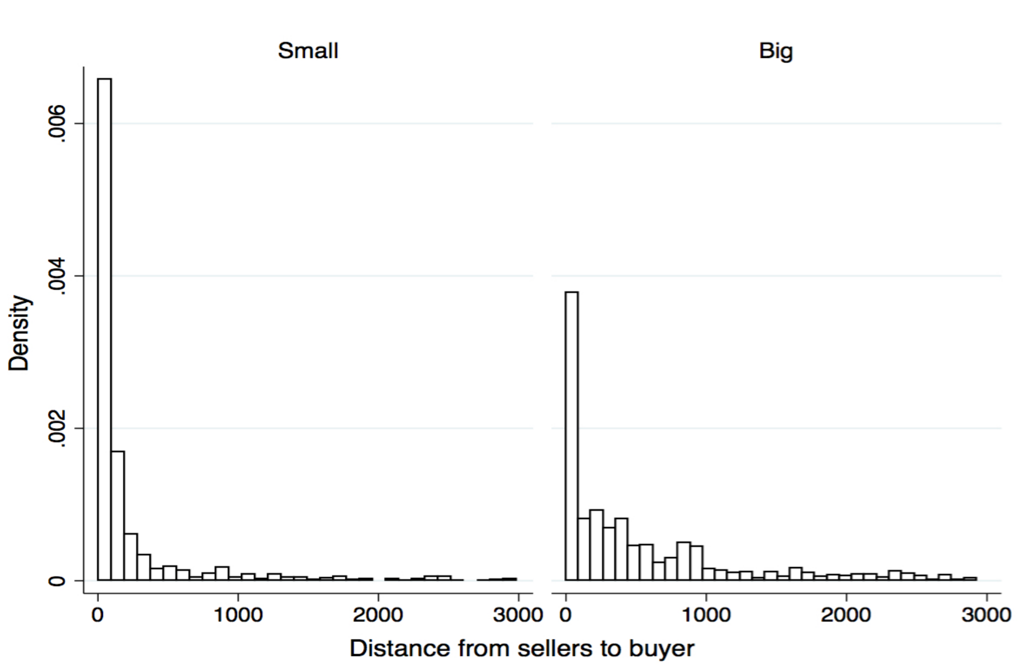



Figure 1.8: Histogram of distances from winners to auction location. Small firms on the left and big firms on the right. Sample consists of off-the-shelf goods.

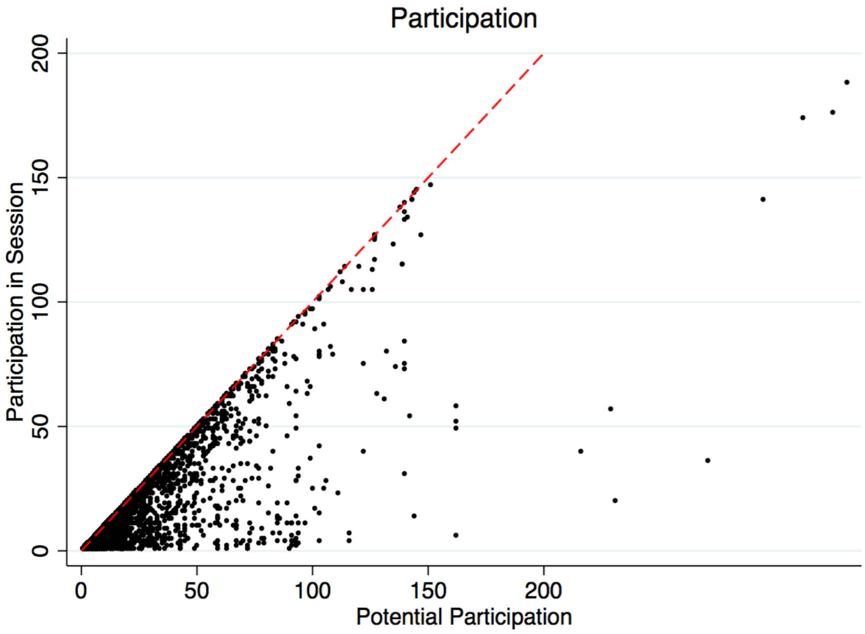



Figure 1.9: Potential participation in a session vs actual participation. Potential participation corresponds to the number of auctions whose product classes belong to the firm portfolio.

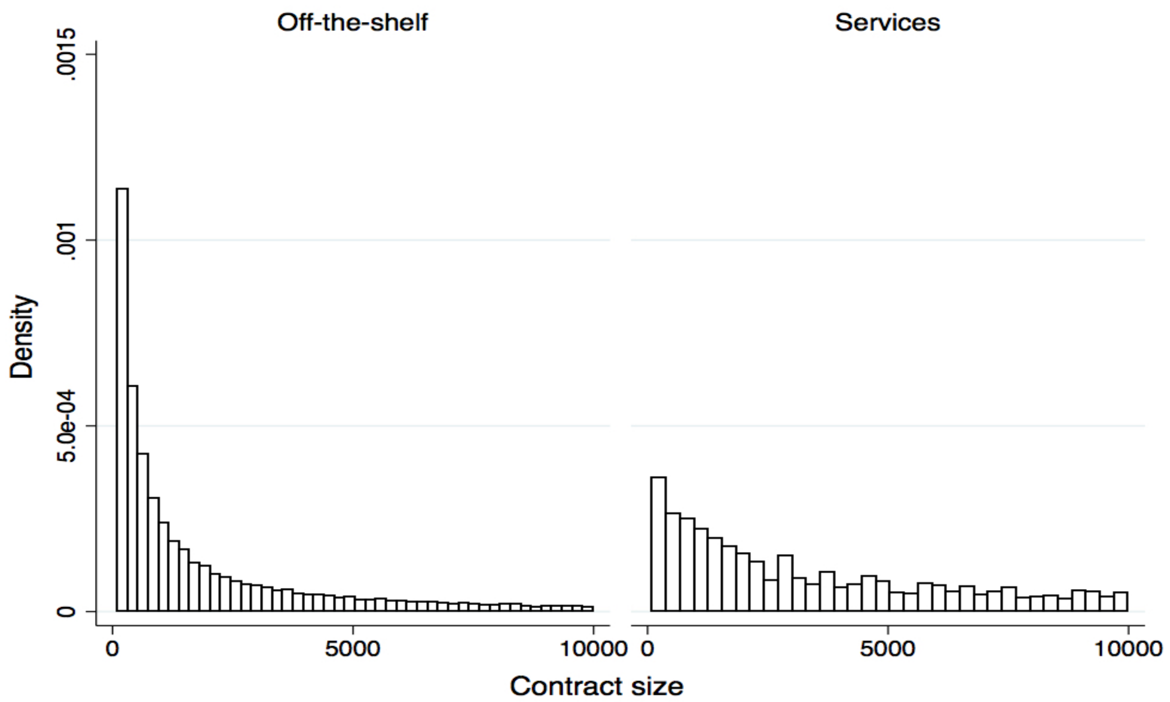



Figure 1.10: Contract size distribution. Histogram truncated at $10,000.

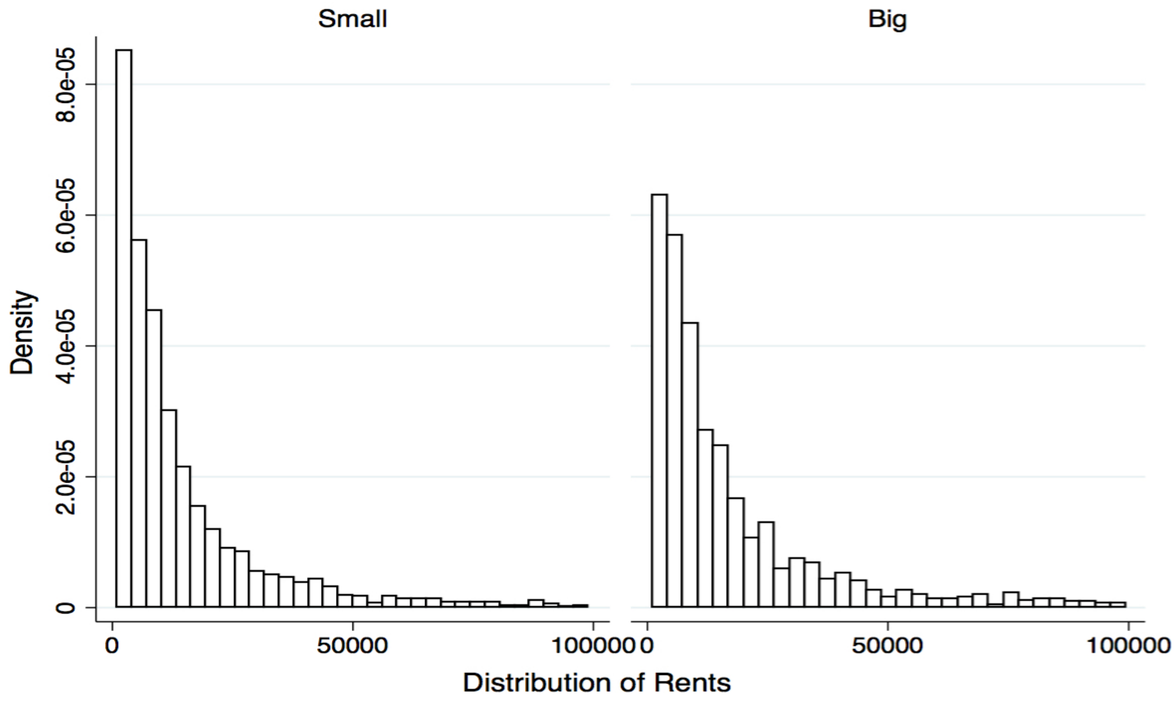



Figure 1.11: Procurement costs distribution allocated to small and big firms. Sample restricted to off-the-shelf contracts with value between \$1,000 and \$10,000.

### Market share by distance

### Equally weighted

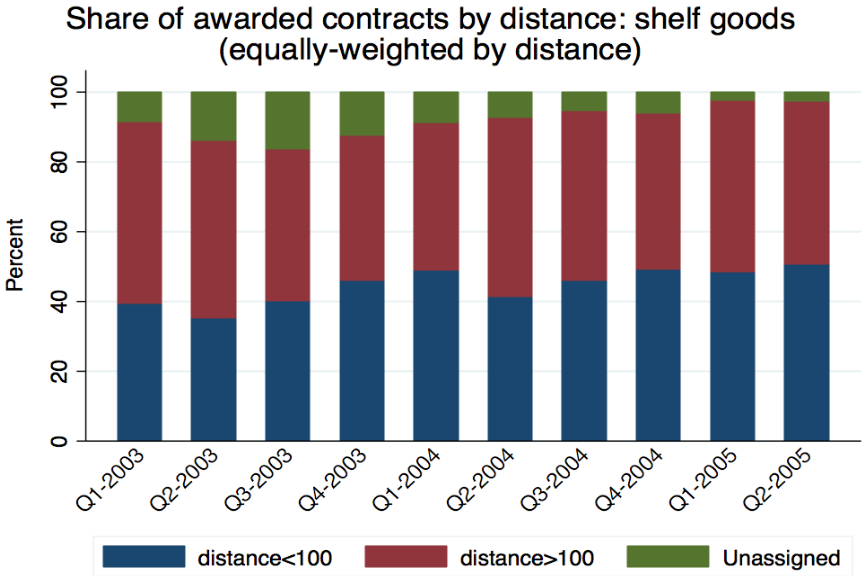



Figure 1.12: Equally-weighted share of contracts by distance to auction location. Sample restricted to 'off the shelf' sessions.

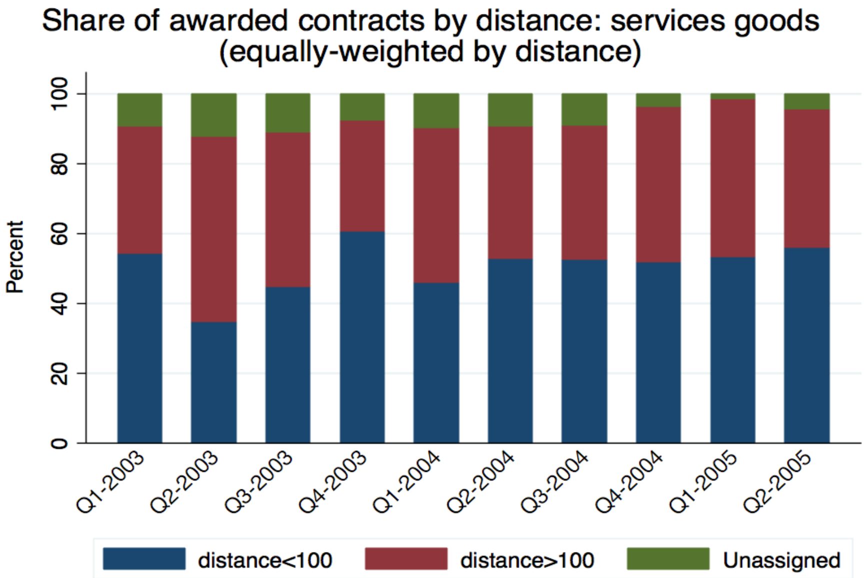



Figure 1.13: Equally-weighted share of 'services' contracts by distance to auction location. Sample restricted to 'services' sessions.

**Value weighted**

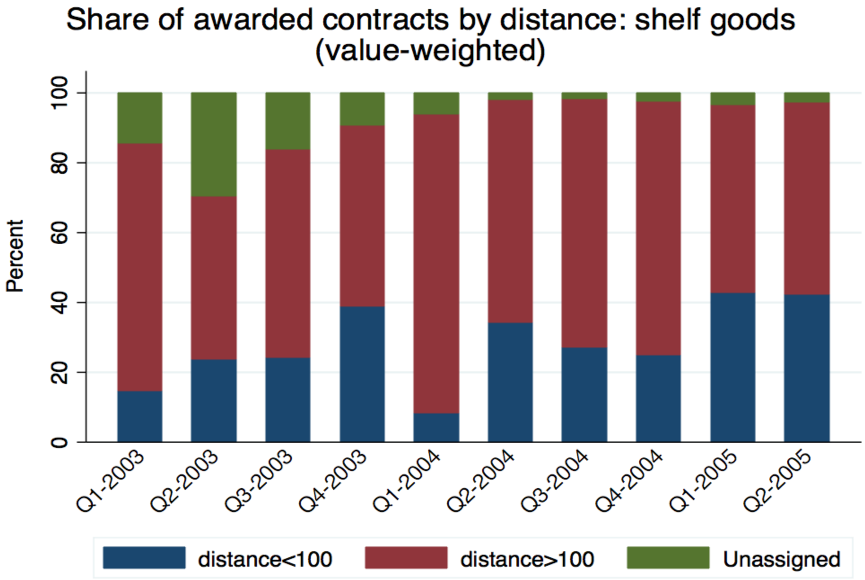



Figure 1.14: Value-weighted share of contracts by distance to auction location. Sample restricted to 'off the shelf' sessions.

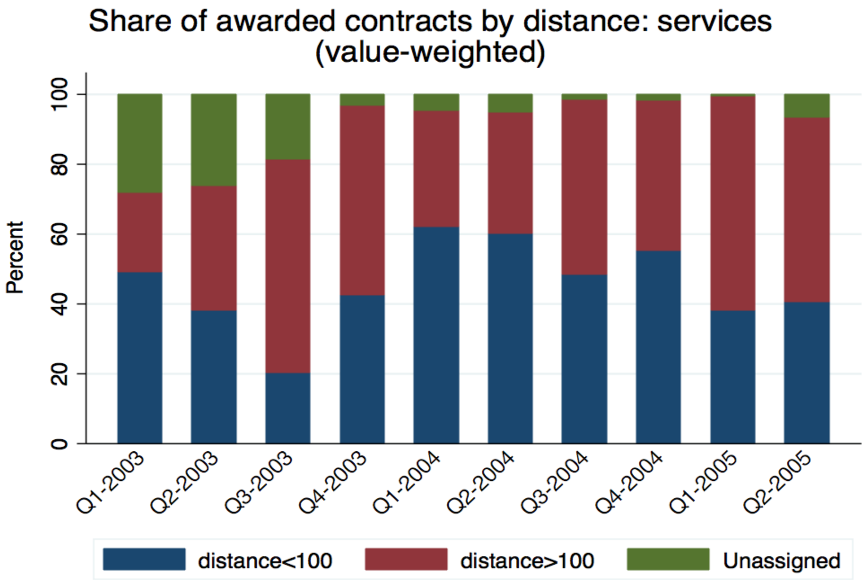



Figure 1.15: Value-weighted share of contracts by distance to auction location. Sample restricted to 'services' sessions.

### Market share by firm size

### Equally weighted

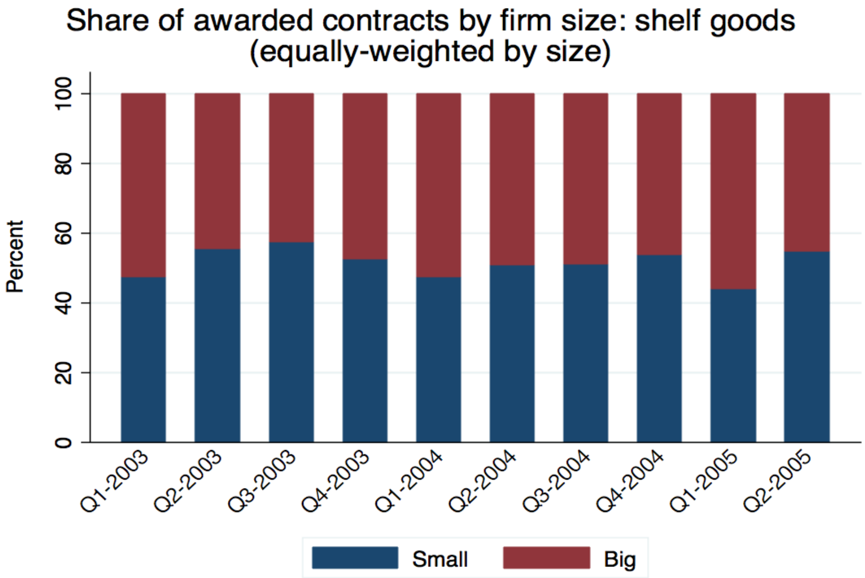



Figure 1.16: Equally-weighted share of contracts by firm size. Sample restricted to 'off the shelf' sessions.

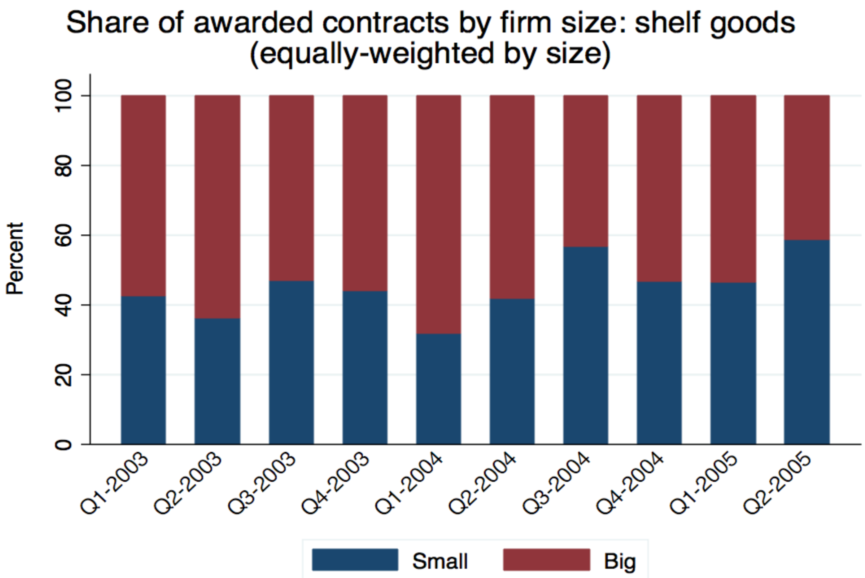



Figure 1.17: Equally-weighted share of contracts by firm size. Sample restricted to 'services' sessions.

**Value weighted**

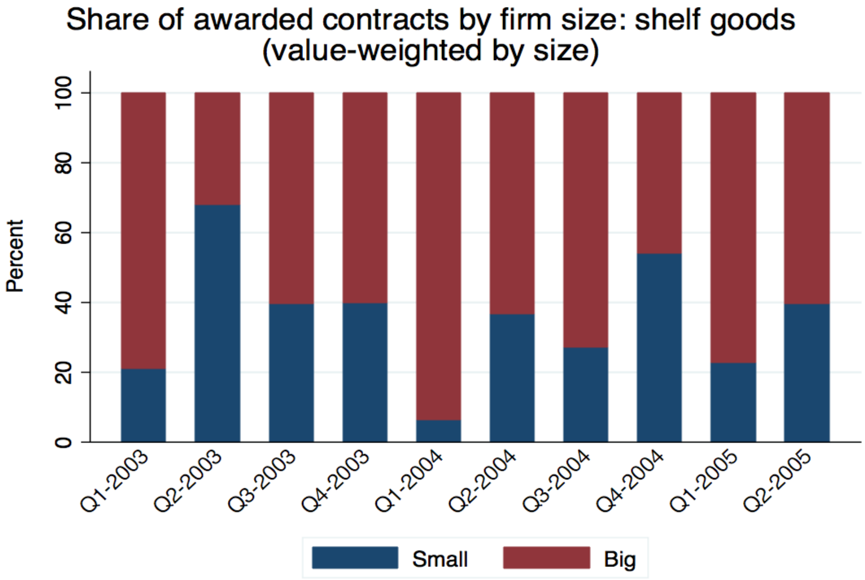



Figure 1.18: Value-weighted share of contracts by firm size. Sample restricted to 'off the shelf' sessions.

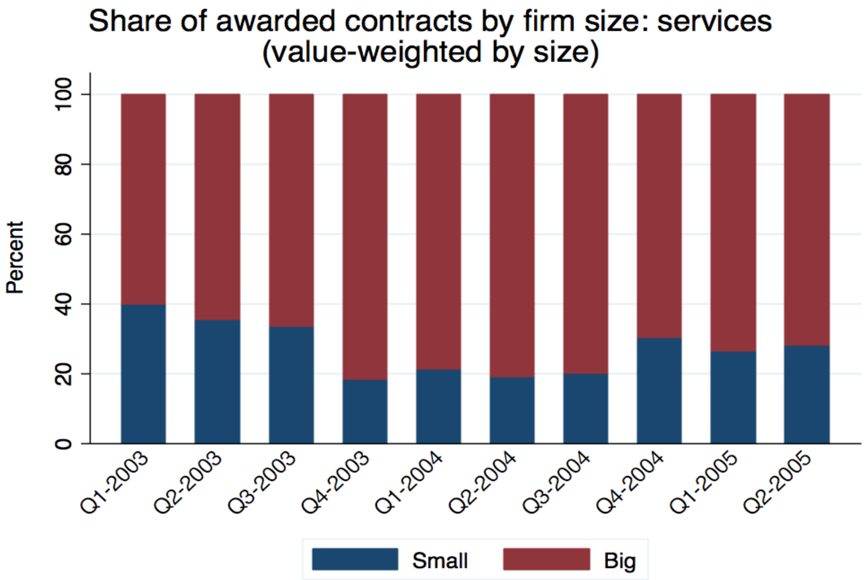



Figure 1.19: Value-weighted share of contracts by firm size. Sample restricted to 'services' sessions.

### Market share by distance and size

### Equally weighted

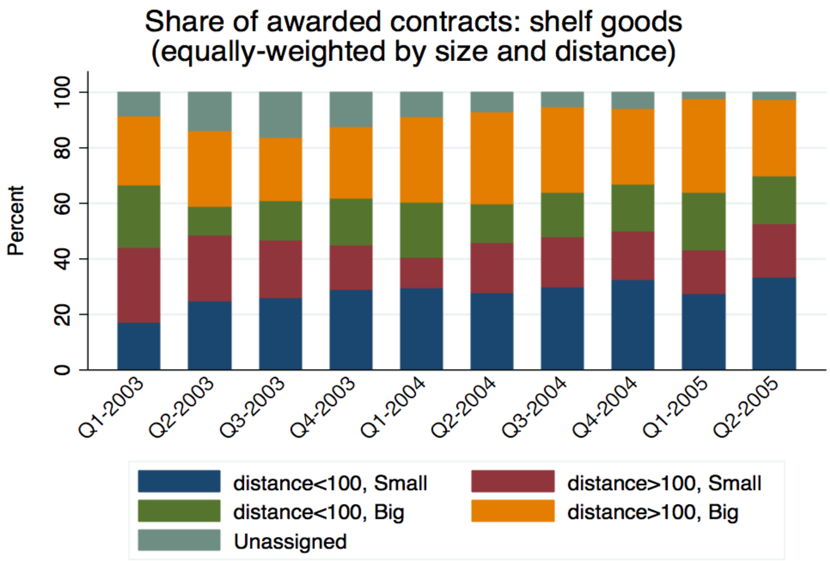



Figure 1.20: Equally-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'off the shelf' sessions.

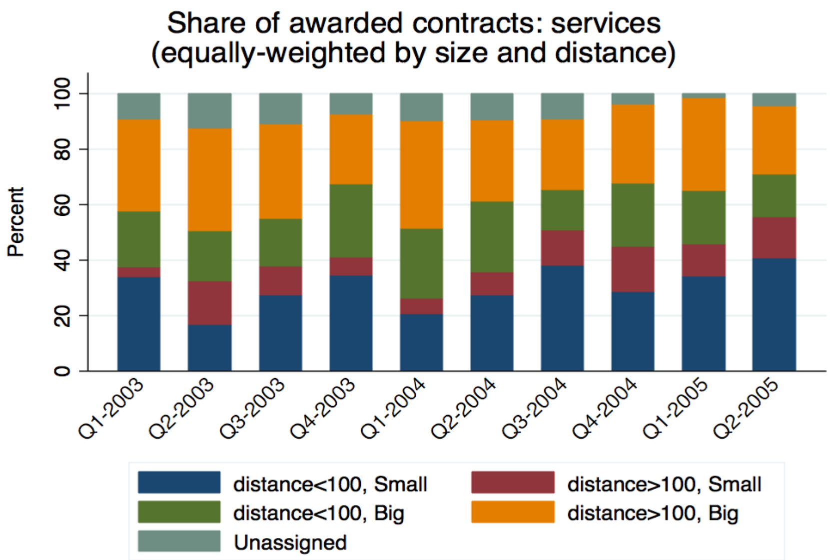



Figure 1.21: Equally-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'services' sessions.

**Value weighted**

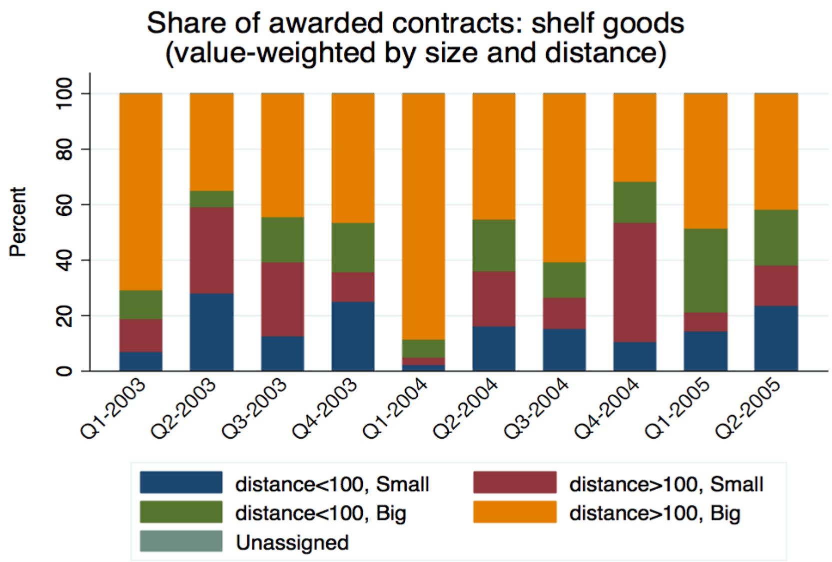



Figure 1.22: Value-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'off the shelf' sessions.

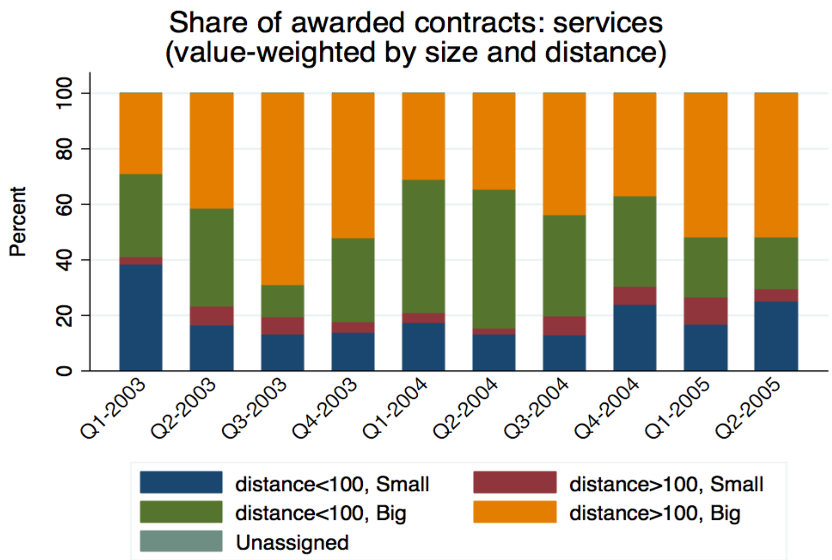



Figure 1.23: Value-weighted share of contracts by distance to auction location and firm size. Sample restricted to 'services' sessions.

# CHAPTER 2

# Menu Choice and Market Structure: An empirical exercise on multi-object auctions

## 2.1 Introduction

The trade-off between efficiency and equity is a central topic in economics. While the concept of efficiency is related to the optimal production and allocation of resources in the economy, equity is related to the distribution of those resources throughout society. Public contracting can be seen through the lens of both concepts. Unambiguously, the government cares about paying the smallest price for a given contract. At the same time, redistribution policies can increase the opportunity to win contracts to a broader set of firms, thereby promoting development across regions. In this chapter, I address the trade-off between efficiency and equity using the case of the Brazilian on-site public procurement sector.

This empirical exercise is motivated by the findings of chapter 1 where I found evidence of high and heterogeneous barriers of entry in this sector. First, I observed that most of the participants are local to the auctions.[1] Second, many firms could have participated in the auctions but chose not to.[2] Third, firms respond differently depending on the design of the section for entry in terms of contract size, location, and product selection.[3] Finally, I showed that procurement costs fall substantially given extra bidder participation.[4] It is an empirical exercise to estimate firms' entry costs in order to analyze the extend to which the selection process can decrease competition and increase procurement costs.

[1]According to the findings of figure 1.11.

[2]Table 1.8.

[3]Tables 1.9 and 1.10.

[4]Table 1.11.

To perform the empirical exercise, I build a novel model of endogenous entry in multi-object auction sessions that allows me to disentangle two forces that affects entry decisions: entry costs and the menu of items of a given session. The model has two stages. In the first stage firms decide whether to enter an auction session and pay a fixed cost after observing an imperfect signal of their true cost. In the second stage, both the items for which they can bid and their costs are realized and the auction takes place. I focus the analysis on type symmetric equilibria, where bidders of the same type follow the same entry strategy. In equilibrium, marginal bidders make zero profits. This condition allows me to link the unobserved entry costs to the observed bid behaviour of entrants.

Having derived the equilibrium of the model, I estimate model fundamentals. The results provide evidence that entry is more attractive to local firms. I find that their cost distribution stochastically dominates the one from non-local firms. Moreover, conditioned on the number of items a firm can participate in, non-local firms face between 2.9% to 7.1% higher entry costs than local firms.

I use the estimated cost distributions and entry costs to simulate two counterfactual scenarios. In the fully efficient scenario, where firms do not incur any entry costs, I find that procurement costs would be lowered by 22.5% to 40.1%. These are bounds on the maximum costs savings and also quantifies the degree of inefficiency present in this market. The second counterfactual is a partially efficient scenario, where non-local firms face the same entry costs as local firms. This analysis focuses on a selected equilibrium where firms enter the sessions sequentially. Firms are sorted according to a lexicographic order which is determined by the strength of their signal, number of items, and firm type (non-local/local). I find that procurement costs would be lowered by 2.8% to 2.9%. Thus, on this type of equilibrium and by holding on-site auctions, the government indirectly sacrificed some efficiency to the benefit of local firms.

As a follow up from chapter 1, the literature review is presented in Section 1.2. The rest of the chapter is as follows. The structural model is presented in section 2.3. Section 2.4 provides the identification results on model fundamentals, while Section 2.5 describes the estimation procedure to recover them. Empirical results are presented in Section 2.6 which includes the counterfactual exercises. Section 2.7 concludes.

## 2.2 Structural Model

A buyer (government) is interested in purchasing items from a listing $\mathcal{A}$. Each item consists of a single unit and is bought individually and simultaneously using an open descending price auction format. There is a set of potential sellers (firms) who may be one of $\tau = 1, ..., \bar{\tau}$ types with $N_\tau$ sellers of type $\tau$.

Sellers have independent private values (IPV) and are endowed with line of products $\mathcal{A}_\tau$. The valuation of a seller of type $\tau$ for item $\omega$ comes from a distribution $G_{\tau,\omega}$, assumed strictly increasing, continuous and differentiable for all types with common support across types $[0, \bar{c}_\omega]$.

Entry into the session is costly. Prior to participating in the session a potential seller of type $\tau$ observes a noisy signal $S_\tau$ and decides whether to enter the session by incurring an entry cost $F_\tau$. The signal is informative on the availability of item $\omega$ in the firm's portfolio and the overall profitability of the session. Entry costs, on the other hand, include several components such as the firms' outside option, costs of related to their distance to the auction location, research costs, and coordination costs from participating in multiple items.

Post entry, firms receive a shock $z_{\tau,\omega} \in \{0, 1\}$ that is informative on the availability of the supply of item $\omega \in \mathcal{A} \cap \mathcal{A}_\tau$.[5] Firms learn their true valuations for each available item and the bidding process takes place. Each individual auction is costless. This is a feature of the data and simplifies the firm problem into a single entry decision: participate or not in the session.

### 2.2.1 Fundamentals

Sellers of type $\tau$ have the following fundamentals:

- Line of products: $\mathcal{A}_\tau$
- Cost distributions: $G_{\tau,\omega}(\cdot)$, $\omega \in \mathcal{A}_\tau$.
- Shock distribution: $f_{z_{\tau,\omega}}(1) = P(z_{\tau,\omega} = 1)$
- Entry costs: $F_\tau$.

[5] This is a restriction to rationalize the empirical observation that firms do not bid for all products in their portfolio. As previously noted in session 1.4.2, firms on average bid for most of their available products.

### 2.2.2 Signal structure

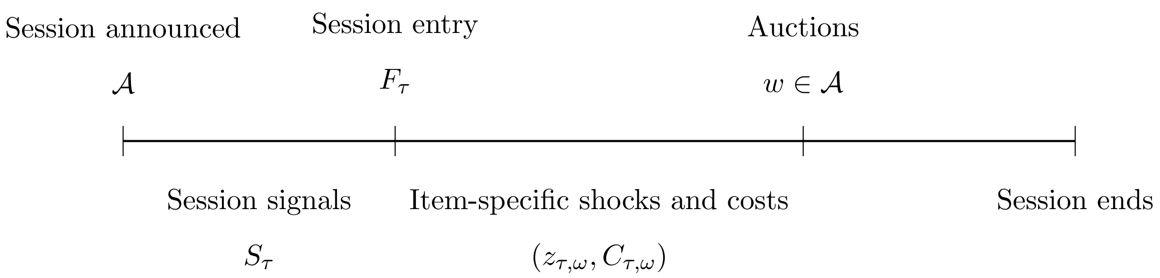
Let $i$ denote a bidder, $\tau$ a type and $\omega$ an item.

- Prior to entry in the session sellers observe a signal on the profitability of the session: $s_{i,\tau}$, where $S_\tau \sim U[0,1]$.
- Post entry bidders observe their true costs $C_{\tau,\omega} \sim G_{\tau,\omega}(\cdot)$.

Having received a signal, a seller forms a posterior belief about his valuation according to Bayes' rule. Let $G_{\tau,\omega}(c_{\tau,\omega}, s_\tau)$ denote the joint cumulative distribution of $(C_{\tau,\omega}, S_\tau)$.

### 2.2.3 Timing

The timing of the two-stage game is as follows:




### 2.2.4 Assumptions

Let $\omega$ be an item, $\tau$ a type, $C_{\tau,\omega}$ denote the random variable describing costs, and $S_\tau$ be the signal observed prior to entry.

**Assumption** (1)**.** *The following structure is imposed to signals and valuations:*

1. $(C_{\tau,1}, ...C_{\tau,\mathcal{A}}, S_\tau)$ *is drawn symmetrically across bidders of type* $\tau$*.*
2. $G_{\tau,\omega}(c,s)$ *is continuous in* $(c,s)$*.*
3. $(S_{i,\tau}, C_{i,\tau,\omega})$ *is independent across bidders of type* $\tau$*.*
4. $E[C_{i,\tau,\omega}|S_{i,\tau} = s]$ *is continuous in* $s$*.*
5. $G_{\tau,\omega}(c|s_\tau)$ *is decreasing in* $s_\tau$*, strictly increasing in* $c$*.*

6. *$C_{\tau,\omega'}$ is conditionally independent from $C_{\tau,\omega}$ given $S_\tau$.*

7. *$C_{\tau,\omega}$ is bounded with support $[0, \bar{c}_\omega]$.*

8. *$S_\tau \sim U[0,1]$.*

Assumptions 1.1-1.4 are standard and Assumption 1.8 is a normalization on signals, without loss of generality. Assumption 1.5 has an economic content and reflects that lower signals are good news, that is, a low signal means a high probability that the seller will be low cost. Assumption 1.6 restricts the informativeness of signals. Once the bidder knows a signal for the session, the cost for one item is not informative for other items.

### 2.2.5 Equilibrium

Following Athey, Levin, and Seira (11) and Roberts and Sweeting (13), I focus on type-symmetric Bayesian Nash equilibria. That is, Nash equilibria where players of the same type use the same strategy. In the open auction stage, bidders know their values so it is weakly dominant for each seller to drop out when the price surpasses their value. In the prior stage of the game, players make entry decisions based on expected profits given their signal. Because signals are independent and low signals are good news, optimal entry decisions involve cutoff signals for each type such that below it a seller would enter the session and above it she would not. This is formalized in lemma (1).

Let the entry decision in the first stage involve entry thresholds, $\bar{s} = (\bar{s}_\tau)_\tau$ such that bidder $i$ of type $\tau$ chooses to enter the auction session if $s_{i\tau} < \bar{s}_\tau$. Let N $= (N_\tau)_\tau$.

The distribution of values among entrants of type $\tau$ for item $\omega$ given entry thresholds is

$$G^*_{\tau,\omega}(c;\bar{s}) = P(C_{\tau,\omega} \leq c | s_\tau \leq \bar{s}_\tau, z_{\tau,\omega} = 1; \bar{s}) \tag{2.1}$$

Denote $G^*_{\tau,\omega,N}(c;\bar{s})$ the probability that an entrant of type $\tau$ with value $c$ for item $\omega$ wins against $\sum_\tau N_\tau - 1$ competitors. An entrant with value $c$ outbids any given potential rival when 1) the rival does not enter; 2) enters and receive a negative shock; 3) enters, receive a positive shock, and has a value above $c$. The probability that this event will take place is derived below:

$$G^*_{\tau,\omega,N}(c;\bar{s}) = f_{z_{\omega,\tau}}(1) \cdot \prod_{\tau'=1}^{\bar{\tau}} [(1-\bar{s}_{\tau'}) + \bar{s}_{\tau'} f_{z_{\omega,\tau'}}(0) + \bar{s}_{\tau'} f_{z_{\omega,\tau'}}(1)(1 - G^*_{\tau',\omega}(c;\bar{s}))]^{N^*_{\tau'}} \tag{2.2}$$

where $N^*_{\tau'} = N_\tau$ if $\tau' \neq \tau$ and $N^*_{\tau'} = N_\tau - 1$ if $\tau' = \tau$.

By the envelope theorem, the expected profit of an entrant of type $\tau$, drawing value $c$ for item $\omega$, competing against $\sum_\tau N_\tau - 1$ potential rivals at entry thresholds is

$$\pi_{\tau,\omega}(c;\bar{s},N) = \int_c^{\bar{c}_\omega} G^*_{\tau,\omega,N}(y;\bar{s})dy \tag{2.3}$$

The ex-ante profit for bidder $i$ of type $\tau$ with signal $s_{i,\tau}$ is then

$$\Pi_{\tau,\omega}(s_{i,\tau};\bar{s},N) = \mathbb{E}[\pi_{\tau,\omega}(c;\bar{s},N)|S_\tau = s_{i,\tau}] = \int_0^{\bar{c}_\omega} G_{\tau,\omega}(y|s_{i,\tau})G^*_{\tau,\omega,N}(y;\bar{s})dy \tag{2.4}$$

Thus, a bidder of type $\tau$ will enter the session if and only if

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{i,\tau};\bar{s},N) \geq F_\tau$$

At the equilibrium entry threshold $s^*_\tau$ firms of type $\tau$ break-even. This gives rise to the following system of equations:

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(\bar{s}^*_\tau;\bar{s}^*,N) = F_\tau \quad \tau = 1,...,\bar{\tau}. \tag{2.5}$$

The following result generalizes Gentry & Li (14) and ensures the existence of a type-symmetric equilibria.

**Proposition 2.2.1.** *The following properties hold:*

*a.* $\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{i,\tau};\bar{s},N)$ *is continuous in* $(s_{i,\tau};\bar{s})$*;*

*b.* $\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{i,\tau};\bar{s},N)$ *is decreasing in* $(s_{i,\tau};\bar{s})$*;*

*c. There exist a type symmetric equilibrium in cutoff strategies such that below it's corresponding cutoff a seller would enter the session and above it she would not. Moreover, any pure strategy* *symmetric equilibria has a payoff equivalent threshold representation.*

*Proof.* See Appendix 2.7

Intuitively, a higher signal is bad news and it decreases expected profits while a higher entry threshold translates to additional competition and also decreases expected profits. While we do not rule out the possibility of multiple equilibria, this is standard in games with entry.[6] It is part of the empirical exercise to deal with this issue.

### 2.2.6 Type space

The reduced form result from session 1.5 motivates the definition of the type space. We found evidence that in multi-object auctions firms face different entry decisions based on the set of feasible set of goods it can sell to the government. This inherently introduces many types to the setting. Moreover, firm location creates asymmetry on the cost distributions that bidders draw from. This suggests that the ideal type space is $\mathcal{T} = (\mathcal{T}_1, \mathcal{T}_2) \in \mathcal{A} \times [0, +\infty)$, where $\mathcal{A}$ is the list of all products that can be purchased.

Such type space imposes a challenge on estimating the model non-parametrically. I proceed with the analysis by making additional assumptions. Namely, I assume that the effect of distance is the same among firms that are either close or far from the session. This assumption allows me to bunch firms according to their distance and simplifies the type to $\mathcal{T} = \mathcal{A} \times \{close, far\}$.[7] I tested three distinct cutoffs for classifying firms according to their distance to the sessions: 100km, 200km, and 300km.

Lastly, the conditional independence assumption 1.6 restricts potential synergies between goods from taking place. This simplifies the effect that the portfolio has on firms costs and allows me to estimate them looking at individual auctions within a session. That is, at the auction level there are only two types: $\{close, far\}$.

[6]See Berry (92) and Tamer (03) for reference.

[7]Types *close* represent local firms and *far* represent non-local firms.

## 2.3 Identification

The objective of identification in this context is to recover informative bounds on cost distributions and entry costs per bidder type.

### Valuation Distribution

Conditional on auction and session observables, heterogeneity in valuations reflects idiosyncratic factors and are assumed to be i.i.d. Following Asker (10) and Haile, Hong & Shum (2003), the valuation of bidder $i$ of type $\tau$ in auction $j$, session $k$ is modeled as follows:

$$c_{i,j,k,\tau} = \Gamma(x_{j,k}) \cdot \tilde{c}_{i,j,k,\tau},\ where\, \Gamma(x_{j,k}) = e^{x_{j,k}\gamma_{j,k}}$$

According to this specification, observed heterogeneity is controlled by projecting log bids onto auction and session characteristics. In order to preserve the observed order statistics, any observable that affects bids symmetrically can enter $x_{j,k}$. This includes, for instance, the type of good and contract specifications but not distance from bidder to session location and quantity supplied which have heterogeneous effects across bidders. This is the primary motivation for including distance in the type space itself. The estimated residuals are assumed to reflect the firms' valuations and are used in the analysis that follows.[8]

The identification of cost distributions follows Haile and Tamer (03). The key idea is to use the mapping between the distribution of order statistics from an i.i.d. sample of size $n$ drawn from a parent distribution $G(\cdot)$ and the distribution itself. If the econometrician observes all order statistics then it is immediate to estimate cost distributions directly from the data. This is not the case here because the proposals stage selects the top bidders to participate in the open descending price auction and not all order statistics are observed. Below, I detail the method. To fix ideas, let us abstract from selection and type space for convenience of notation:

[8] I use the estimates from the Heckman two-step regression for the contract quantity as it accounts for selection. The estimate from the third column of table 1.10 was chosen as it reflects detail product code: $\gamma_q = 0.68$.

$$G^{i:n}(c) = \frac{n}{(n-i)! \cdot (i-1)!} \int_0^{G(c)} t^{i-1}(1-t)^{n-i} dt \tag{2.6}$$

Costs distributions are assumed strictly increasing (Assumption 1.5). This guarantees that equation (6) uniquely defines the quantiles of $G$. Let $\phi$ the inverse functional such that $\phi(G^{i:n}(c); i, n) = G(c)$. Thus, in order to identify a cost distribution $G(\cdot)$ it is sufficient that one identify its order statistics.

Unfortunately, we cannot point identify order statistics directly from the data. This is because the auction follows an open descending price format and, as such, there are jump bids. Nevertheless, as long as bidders undercut each other when it is profitable to do so (Assumption 3a) and bid above their costs (Assumption 3b), drop out points induce informative bounds on bidders' values.[9] That is, once a bidder places a bid $b_1$ and drop out facing the outstanding bid is $b_2$ we can infer that a bidder's cost $c$ satisfy $b_2 < c < b_1$.[10]

**Assumption** (3)**.** *The following behavioural assumptions hold:*

*3a. Bidders do not let someone win at a price they are willing to beat*

*3b. Bidders do not bid less than they want to be paid*

From these bounds and the operator $\phi$ we can induce a stochastic ordering between order statistics of cost distributions and the parent distribution. Let $U^{i,n}$ denote the random variable describing the last bid submitted by a bidder prior to its drop out point that ranks $i$ out of $n$ bidders.

$$\begin{aligned} U^{i,n} \geq c^{i:n} &\Rightarrow G_U^{i,n}(c) \leq G^{i:n}(c) \\ &\Rightarrow \phi(G_U^{i,n}(c); i, n) \leq G(c) \\ &\Rightarrow \max_{i,n}\{\phi(G_U^{i,n}(c); i, n) \leq G(c) \end{aligned}$$

[9]In a clock auction, prices decrease continuously and there are no jump bids. This allows for point identification of order statistics.

[10]Recall that bidders are approached sequentially by the auctioneer and have to make a decision whether to undercut the outstanding bid or drop out.

Let $L_i^{i,n}$ be the outstanding bid at which this bidder drops out. The only bidder who does not have a drop out point from the auction is the winner (i = 1).

$$
\begin{aligned}
c^{i:n} \geq L_i^{i,n},\, i \geq 2 \Rightarrow G^{i:n}(c) &\leq G_L^{i,n}(c),\, i \geq 2 \\
\Rightarrow G(c) &\leq \phi(G_L^{i,n}(c); i, n),\, i \geq 2 \\
\Rightarrow G(c) &\leq \min_{i \geq 2, n} \{\phi(G_L^{i,n}(c); i, n)\}
\end{aligned}
$$

As long as bid increments are small, these bounds are informative. In fact, our setting allows us to improve on the lower bounds of Haile and Tamer.[11] This follows from the design of our setting where the auctioneer asks bidders sequentially if their are willing to undercut the outstanding bid. A negative answer automatically drops them out of the auctions and this is recorded in the auction minutes. It is important to point out that this is unusual in open descending price auctions and it is often the case that drop out points are not identified. In this case, the only lower bound that is informative is the winning bid which, correcting for bid increments, serves as a lower bound for the second order statistic. Nevertheless, the bounds obtained here are still not sharp as we do not use the joint distribution of order statistics to recover cost distributions.[12]

The estimator for these lower and upper bounds are easily computed through their corresponding sample analogs. However, they might cross in small samples. This is because the sample analog for the upper (lower) bound is downward (upward) biased. One option around this issue is to follow Haile and Tamer (03) who proposes a smoothed version of the estimators with a parameter that is used to guarantee convergence and no crossing. Their estimator is as follows:

$$
f(y_1, ...y_n; \rho) = \sum_{i=1}^{n} y_i \Big[ \frac{e^{y_i \rho}}{\sum_{k=1}^{n} e^{y_i \rho}} \Big],
$$

where $f(y_1, ...y_n; \rho) \xrightarrow{\rho \to -\infty} \min\{y_1, ..., y_n\}$ and $f(y_1, ...y_n; \rho) \xrightarrow{\rho \to \infty} \max\{y_1, ..., y_n\}$.

A caveat for this approach comes from the fact that these estimators are constructed without incorporating a measure of sample errors. However, their closed-form solution makes them practical

[11] upper bounds in the case of procurement.

[12] For a further discussion on this topic see Haile and Tamer (03), Appendix D.

and convenient for empirical implementation.[13]

Having derived appropriate bounds for cost distributions, bounds on entry cost are immediately induced from the zero profit condition of marginal bidders, equation (2.5).

**Lemma 2.3.1.** *Given assumptions 1-3, bounds on cost distributions induce bounds on entry costs.*

## 2.4 Estimation Procedure

The zero profit condition of marginal bidders (equation 2.5) is key for the estimation of entry costs. It separates entrants from non-entrants and creates natural bounds for entry costs:

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(\bar{s}^*_\tau; \bar{s}^*, N) < F_\tau \quad if \ s_{i,\tau} \leq \bar{s}_\tau$$

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(\bar{s}^*_\tau; \bar{s}^*, N) > F_\tau \quad if \ s_{i,\tau} > \bar{s}_\tau$$

Bidder signals are not observed and, as such, the two conditions above are suitable for the estimation strategy. Below, I detail the moments inequalities approach to estimate bounds on entry costs.

Let n be the number of bidders in a given auction, $i$ be an order statistic and $\hat{s}_\tau$ be the proportion of entrants of type $\tau$ into the session. Estimators for the bounds on cost distributions are constructed as follows:[14]

---

[13] A second option that incorporates a measure of sample errors is Chernokhukov, Lee, and Rosen (14) who propose a precision-corrected estimate for the lower and upper bounds. Their method is as follows:

$$\hat{G}^{i,n}_U(c,p) := min_{(i,n), i \geq 2}\Big\{\phi(\hat{G}^{i,n}_U(c); i, n) + k(c,p) \cdot s_U(i,n)\Big\}$$

$$\hat{G}^{i,n}_L(c,p) := max_{(i,n)}\Big\{\phi(\hat{G}^{i,n}_L(c); i, n) - k(c,p) \cdot s_L(i,n)\Big\}$$

where $s_U(i,n)$ is the standard error of $\phi(\hat{G}_U(c); i, n)$ and $k(c,p)$ is a critical value derived from an appropriate sequence of Gaussian processes.

[14] I focus on the sample of auctions where all firms had the same type in order to avoid estimating mixing distributions.

$$G_{U,\tau,\omega,i,n}(c;\hat{s}) = \frac{1}{T_n}\sum_{t=1}^{T}\mathbb{1}\Big\{n_t = n, x = i, s \leq \hat{s}, U \leq c\Big\}$$

$$G_{L,\tau,\omega,i,n}(c;\hat{s}) = \frac{1}{T_n}\sum_{t=1}^{T}\mathbb{1}\Big\{n_t = n, x = i, s \leq \hat{s}, L \geq c\Big\}$$

$$T_n = \sum_{t=1}^{T}\mathbb{1}\Big\{n_t = n, x = i, s \leq \hat{s}\Big\}$$[15]

1. Estimate entry frequencies $\hat{s}_\tau = \frac{E[n_\tau|N_\tau]}{N_\tau}$, where $N_\tau$ is the set of potential bidders of type $\tau$ and $n_\tau$ denote the number of entrants;

2. Estimate conditional order statistics $\hat{G}_{U,\tau,\omega,i,N}(c|s \leq \hat{s}_\tau; \hat{s}, z_{\tau,\omega} = 1)$ and $\hat{G}_{L,\tau,\omega,i,N}(c|s \leq \hat{s}_\tau; \hat{s}, z_{\tau,\omega} = 1)$;

3. Estimate conditional cost distributions $\hat{G}_{U,\tau,\omega}(c|s \leq \hat{s}_\tau; \hat{s}, z_{\tau,\omega} = 1)$ and $\hat{G}_{L,\tau,\omega}(c|s \leq \hat{s}_\tau; \hat{s}, z_{\tau,\omega} = 1)$. Retrieve 'best' lower and upper bounds.

4. Estimate shock distribution $\hat{f}_{z_\tau}(1) = \frac{\#\mathcal{A}^*_\tau}{\#\mathcal{A}\cap\mathcal{A}_\tau}$, where $\mathcal{A}^*_\tau$ is the observed set of items a firm bid for and $\#\mathcal{A}\cap\mathcal{A}_\tau$ is the set of feasible items the firm could bid for;

5. Estimate bounds on winning probabilities

$$\hat{G}^*_{U,\tau,\omega,N}(c;\hat{s}) = f_{z_{\omega,\tau}}(1)\cdot\prod_{\tau'=1}^{\bar{\tau}}[(1-\hat{s}_{\tau'}) + \hat{s}_{\tau'}\hat{f}_{z_{\omega,\tau'}}(0) + \hat{s}_{\tau'}\hat{f}_{z_{\omega,\tau'}}(1)(1-\hat{G}^*_{L,\tau',\omega}(c;\hat{s}))]^{N^*_{\tau'}}$$

$$\hat{G}^*_{L,\tau,\omega,N}(c;\hat{s}) = f_{z_{\omega,\tau}}(1)\cdot\prod_{\tau'=1}^{\bar{\tau}}[(1-\hat{s}_{\tau'}) + \hat{s}_{\tau'}\hat{f}_{z_{\omega,\tau'}}(0) + \hat{s}_{\tau'}\hat{f}_{z_{\omega,\tau'}}(1)(1-\hat{G}^*_{U,\tau',\omega}(c;\hat{s}))]^{N^*_{\tau'}}$$

where $N^*_{\tau'} = N_\tau$ if $\tau' \neq \tau$ and $N^*_{\tau'} = N_\tau - 1$ if $\tau' = \tau$.

6. Estimate bounds on ex-ante variable profits at the auction level for entrants:[16]

[15] Haile and Tamer (02,03) derive the uniform consistency, asymptotic distribution and bootstrap confidence bands of these estimators. Each conditional cumulative distribution is estimated with at least 15 points.

[16] I use a rectangle integration to approximate the integral.

$$\Pi^{entrant}_{U,\tau,\omega}(s_{i,\tau};\bar{s},N)=\int_0^{\bar{c}_\omega} G_{\tau,\omega}(y|s_{i,\tau})G^*_{U,\tau,\omega,N}(y;\bar{s})dy$$

$$\Pi^{entrant}_{L,\tau,\omega}(s_{i,\tau};\bar{s},N)=\int_0^{\bar{c}_\omega} G_{\tau,\omega}(y|s_{i,\tau})G^*_{L,\tau,\omega,N}(y;\bar{s})dy$$

7. Use variation on entry to estimate bounds on ex-ante variable profits at the auction level for non-entrants had they entered:

$$\Pi^{non\,entrant}_{U,\tau,\omega}(s_{i,\tau};\bar{s}+\epsilon,N)=\int_0^{\bar{c}_\omega} G_{\tau,\omega}(y|s_{i,\tau})G^*_{U,\tau,\omega,N}(y;\bar{s}+\epsilon)dy$$

$$\Pi^{non\,entrant}_{L,\tau,\omega}(s_{i,\tau};\bar{s}+\epsilon,N)=\int_0^{\bar{c}_\omega} G_{\tau,\omega}(y|s_{i,\tau})G^*_{L,\tau,\omega,N}(y;\bar{s}+\epsilon)dy$$

where $\epsilon$ is the effect that a non-entrant bidder would have on the entry frequency.

8. Estimate average ex-ante variable profits at the session level for entrants (non-entrants):

$$\overline{\Pi}^{entrant}_{U,\tau}=\sum_{\omega\in\mathcal{A}_\tau}\overline{\Pi}^{entrant}_{U,\tau,\omega}(\bar{s},N)$$

$$\overline{\Pi}^{entrant}_{L,\tau}=\sum_{\omega\in\mathcal{A}_\tau}\overline{\Pi}^{entrant}_{L,\tau,\omega}(\bar{s},N)$$

9. Construct moments.

Here, I specify a functional form for entry costs:: $F_\tau=\left(\#\mathcal{A}^*_\tau\right)^\alpha\times(F_1+\mathbb{1}_\tau\cdot F_2)$.

$$v_{1,\tau}=\left(-\overline{\Pi}^{entrant}_{U,\tau}+\left(\#\mathcal{A}^*_\tau\right)^\alpha\times\cdot(F^u_1+\mathbb{1}_\tau\cdot F^u_2)\right)^+$$

$$v_{2,\tau}=\left(-\overline{\Pi}^{entrant}_{L,\tau}+\left(\#\mathcal{A}^*_\tau\right)^\alpha\times\cdot(F^l_1+\mathbb{1}_\tau\cdot F^l_2)\right)^+$$

$$v_{3,\tau} = \left( \overline{\Pi}_{U,\tau}^{non-entrant} - \left(\#\mathcal{A}_\tau^*\right)^\alpha \times \cdot (F_1^u + \mathbb{1}_\tau \cdot F_2^u) \right)^+$$

$$v_{4,\tau} = \left( \overline{\Pi}_{L,\tau}^{non-entrant} - \left(\#\mathcal{A}_\tau^*\right)^\alpha \times \cdot (F_1^l + \mathbb{1}_\tau \cdot F_2^l) \right)^+$$

where $X^+ = max(0, X)$.

10. Minimize the criterion function $Q(F) = \sum_{j,\tau} v_{j,\tau}^2$.

I report the results from each step of the estimation procedure in Appendix II.

## 2.5 Empirical Results

I performed the estimation method detailed in the previous session with different cutoffs for defining *close* and *far* firms (100km, 200km, and 300km). I also used the samples of auctions where the number of players were less (or equal) to three, and less than ten to mitigate the potential effects of the two-step stage design. Finally, I focused on the auctions where all firms were either *close* or *far* in order to avoid estimating mixing distributions. The estimated truncated cost distributions are similar, and so I report them focusing on the 300km cut-off for defining firm types. I did not find major differences estimating cost distributions with up to three or ten bidders participating.

First, the empirical distributions on order statistics, reported in figure 2.1, behave well, as there are only a few instances where the lower and upper bounds cross. Second, they confirm the results from the Heckman estimates where I found that larger distances increase firms' variable costs. This translates into firms with type *close* stochastically dominating firms with type *far*. Figure 2.2 is a typical example of that. Third, bounds on the cost distribution are recovered by applying equation 2.6 to the empirical bounds on order statistics, generating figure 2.3. Finally, the 'best' upper and lower bounds are computed according to Haile & Tamer's method for smoothing the minimum operator on upper bounds and maximum operators on the lower bounds.[17] This results in figure 2.3 which shows that the gap between the bounds according to firm type is relatively small.

[17] I use $\rho = log(\sqrt{(\#Sessions)} = 5.4296$ as the smoothing parameter.

The last piece to be obtained prior to estimating entry costs are the shocks players receive post entry. This is to rationalize the fact that many players choose not to bid on an item even though it belongs to their portfolio.[18] Figure 2.6 displays the scatter plot between the number of auctions in a firm's feasible set and the fraction of auctions it actually participated in. Here, a firm's feasible set is defined as the intersection between the firm's line of products and the menu of items auctioned at the session level. This figure suggests that the shock approach is reasonable as there is no correlation between the number of auctions a firm can bid for in a session and the fraction of auctions it actually bid for. I estimate that a firm has, on average, a 81.5% chance of bidding for an item in their feasible set.

After recovering the cost distributions - conditional on entry, I follow the estimation steps 5-9 to recover firms' profits for both entrants and non-entrants.[19] The mean session profits are reported in table 2.1. In general, this table shows that entrants had more items to bid for and a higher variable profit per item. For example, focusing on the 300km cutoff for defining firm type, *close* entrants had approximately 7 items to bid for while non-entrants of either type had 5. At the same time, the *close* non-entrants would have made 19.1% less variable profits per item than the corresponding entrants.

Finally, the set of moment inequalities is constructed using the conditions that firms enter the session if and only if it is profitable to do so. By minimizing the square loss of these moments I am able to estimate bounds on entry costs per item which are reported in table 2.2. From them we are able to takeaway some important points.

First, entry costs are high in order to keep potential bidders out. Besides capturing preparation costs for each item, the component $F_1$ also include the bidders' outside option. It's magnitude provide evidence that the outside option is the main force keeping potential bidders out and it works as a proxy for a player's opportunity cost of participating in a session. Second, conditioned on the number of items a firm can participate in, firms who are *far* face between 2.9% to 7.1% higher entry costs than firms who are *close*. Third, entry costs have the property of decreasing returns to scale regarding the number of items a firm can participate in ($\alpha < 1$). That is, the entry cost per item of a firm decreases as it has a higher session match. However, it is relevant

[18] For example, firms might learn post entry in the session that they would not be able to fulfil a contract.

[19] The latter is possible as the sessions present enough variation on entry

to point out that the current model does not allow firms to select the number of items it would participate in. This would create a double selection problem of entry in sessions and into a subset of auctions that is intractable due to its combinatorial nature. Here, $\alpha$ is not partially identified from moments related to the subsets of items a firm participates in. Rather, its bounds come from moments relating firms that entered a session and the ones that did not. All in all, entry is more attractive to local firms.

Having these estimates in hand, I turn to two policy questions. I focus on quantifying the degree of inefficiency present in the sessions in terms of procurement cost. To do so, I simulate two scenarios.[20]

In the fully efficient scenario, bidders do not face any fixed costs and entry is weakly optimal ($F_1 = 0$ and $F_2 = 0$). By assuming that all potential bidders enter each session I find that, on average, procurement costs would be reduced by 22.5% to 40.1%.[21] These constitutes bounds on the maximum costs savings and quantifies the degree of inefficiency present in this market.

In the partially efficient scenario, bidders have homogeneous entry costs conditional on the number of auctions it would like to participate in ($F_2 = 0$). Nevertheless, bidders remain heterogeneous in three dimensions: signals, portfolio, and cost distribution. This heterogeneity, in addition to the existence of multiple equilibria, makes the counterfactual intractable. To make progress, I restrict the type of equilibrium being played. I assume that bidders enter sequentially with an lexicographic order respectively determined by 1) their signal; 2) number of items; 3) local/non-local firm. While this puts several restrictions on the equilibrium the will be played in the simulation, it is still compatible with the moment inequalities approach for estimating entry costs.[22] Moreover, the analysis focus on the comparison between the equilibrium from the counterfactual scenario and the *same* type of equilibrium generated by the model parameters. I find that procurement costs would be lowered by 2.8% to 2.9%.[23] Thus, on this type of equilibrium and by holding on-site

[20] Simulations used the 300km cut-off to define local firms.

[21] Bootstrap procurement costs for auctions with more than one participant. Re-sampling of 50 times, each time drawing 1000 sessions. Standard deviation of lower bound is 4.7% and standard deviation of upper bound is 3.3%.

[22] It is still complex to check for equilibrium conditions in this setting, as an entrant does necessarily participate in the same set of auctions as the previous bidders.

[23] Bootstrap values for auctions with more than one participant. Re-sampling of 100 times, each time drawing 100 sessions.

auctions, the government indirectly sacrificed some efficiency to the benefit of local firms.

## 2.6 Conclusion

In this chapter I studied the trade-off between efficiency and equity in the Brazilian on-site public procurement sector. I built a novel model of endogenous entry in multi-object auctions which allowed me to disentangle the effects that entry costs and the menu of products have on firms' entry decisions. I showed that this market is far from efficient and that, in a sequential equilibrium, the government forgone some efficiency and indirectly subsidized local firms by holding on-site sessions. This is a consequence of local firms benefiting from the design of the session by having lower costs of entry compared to non-local firms. It is ongoing work to use the empirical results to explore the potential effects of redistributing auctions across different sessions in a given buyer location.

This project opens a research agenda for new theoretical and empirical developments in multi-object auctions. A current limitation of the model is the restriction on the number of items a firm can participate in within a session. Relaxing this restriction will allow the model to generate a richer set of moment inequalities based on revealed preference. For example, one direction of research is to explore moments relating the choice of auctions in a session with any other feasible subset of auctions within the same session. In terms of the empirical framework, continued research is needed to understand the role that unobserved heterogeneity could play in estimating both the valuation distribution and entry costs. It would also be interesting to study the existence of potential synergies from participation in multiple auctions.

The case of Brazil can be used as a benchmark for other studies. On-site sessions are still a common practice and it is common to observe sessions with multiple items being auctioned. The distinction relies on Brazil's exceptional data quality, with detailed information about each bid, item, and firm characteristics. It is the role of future research to continue exploring this data.

## 2.7 Appendix I: Proofs

### 2.7.1 Proof of Proposition 2.2.1

The proofs below extend Gentry and Li (14).

*Proof.* a. Let $(s, \bar{s})$ and $(s', \bar{s}')$ be two pairs of signals with $||(s, \bar{s}) - (s', \bar{s}')|| \to 0$.

$$\left| \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s; \bar{s}, N) - \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s'; \bar{s}', N) \right| \leq$$
$$\leq \sum_{\omega \in \mathcal{A}_\tau} \left| \Pi_{\tau,\omega}(s; \bar{s}, N) - \Pi_{\tau,\omega}(s'; \bar{s}', N) \right|$$

$\Pi_{\tau,\omega}(s; \bar{s}, N)$ is continuous in $(s; \bar{s})$:

$$\left| \Pi_{\tau,\omega}(s; \bar{s}, N) - \Pi_{\tau,\omega}(s'; \bar{s}', N) \right| \leq \tag{2.7}$$
$$\leq \left| \Pi_{\tau,\omega}(s; \bar{s}, N) - \Pi_{\tau,\omega}(s; \bar{s}', N) \right| + \left| \Pi_{\tau,\omega}(s; \bar{s}', N) - \Pi_{\tau,\omega}(s'; \bar{s}', N) \right| \leq \tag{2.8}$$
$$\leq \int_0^{\bar{c}_\omega} \left| G^*_{\tau,\omega,N}(y; \bar{s}) - G^*_{\tau,\omega,N}(y; \bar{s}') \right| dy + \int_0^{\bar{c}_\omega} \left| G_{\tau,\omega}(y|s) - G_{\tau,\omega}(y|s') \right| dy \leq \tag{2.9}$$
$$\leq \int_0^{\bar{c}_\omega} \left| G^*_{\tau,\omega,N}(y; \bar{s}) - G^*_{\tau,\omega,N}(y; \bar{s}') \right| dy + \left| \int_0^{\bar{c}_\omega} \Big( G_{\tau,\omega}(y|s) - G_{\tau,\omega}(y|s')] \Big) dy \right| = \tag{2.10}$$
$$= \int_0^{\bar{c}_\omega} \left| G^*_{\tau,\omega,N}(y; \bar{s}) - G^*_{\tau,\omega,N}(y; \bar{s}') \right| dy + \left| E[C_{\tau,\omega}|S = s] - E[C_{\tau,\omega}|S = s'] \right| \tag{2.11}$$

where (9) comes from stochastic ordering. Also,

$$G^*_{\tau,\omega,N}(c; \bar{s}) = f_{z_{\omega,\tau}}(1) \cdot \prod_{\tau'=1}^{\bar{\tau}} [(1 - \bar{s}_{\tau'}) + \bar{s}_{\tau'} f_{z_{\omega,\tau'}}(0) + \bar{s}_{\tau'} f_{z_{\omega,\tau'}}(1)(1 - G_{\tau',\omega}(c, \bar{s}_{\tau'}))]^{N^*_{\tau'}}$$

Thus, $G^*_{\tau,\omega,N}(c; \bar{s})$ is continuous in $\bar{s}_\tau$ by the continuity of $G_{\tau,\omega}(c, s)$ and

$$\left| \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s; \bar{s}, N) - \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s'; \bar{s}', N) \right| \leq$$

$$\leq \sum_{\omega \in \mathcal{A}_\tau} \int_0^{\bar{c}_\omega} \left| G^*_{\tau,\omega,N}(y; \bar{s}) - G^*_{\tau,\omega,N}(y; \bar{s}') \right| dy$$

$$+ \sum_{\omega \in \mathcal{A}_\tau} \left| E[C_{\tau,\omega} | S = s] - E[C_{\tau,\omega} | S = s'] \right| \to 0$$

b. From

$$\bar{s}_\tau G^*_{\tau,\omega,N}(c; \bar{s}) = P(C_{\tau,\omega} \leq c, s_\tau \leq \bar{s}_\tau) = \int_0^{\bar{s}_\tau} G_{\tau,\omega}(c|s) ds$$

We get that the winning probability is decreasing at the cutoffs ($c \neq 0$)

$$\frac{\partial G^*_{\tau,\omega,N}(c; \bar{s})}{\partial \bar{s}_\tau} = f_{z_{\omega,\tau}}(1) \cdot \prod_{\tau' \neq \tau}^{\bar{\tau}} [(1 - \bar{s}_{\tau'}) + \bar{s}_{\tau'} f_{z_{\omega,\tau'}}(0) + \bar{s}_{\tau'} f_{z_{\omega,\tau'}}(1)(1 - G^*_{\tau',\omega}(c; \bar{s}))]^{N^*_{\tau'}} \times$$
$$\times \Big( -1 + f_{z_{\omega,\tau}}(0) + f_{z_{\omega,\tau}}(1)(1 - G_{\tau,\omega}(c|\bar{s}_\tau) \Big) \leq 0$$

Also, from stochastic ordering $G_{\tau,\omega}(y|s)$ is decreasing in $s$. From (4) it follows that expected session profits are decreasing in $(s_{i,\tau}; \bar{s})$.

c. First, note that $\sum_{\omega \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{i,\tau}; \bar{s}, N)$ is decreasing in $s_{i,\tau}$. That is, higher signals are bad news. Moreover, since a player's signal is only informative on his own costs, a cutoff strategy with entry below a certain threshold is a best-response.

Let

$$\Xi_1 = \{\tau \in \bar{\tau} | \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(\bar{s}_\tau; (\bar{s}_\tau, \bar{s}_{-\tau}) = (\bar{s}_\tau, 1), N) > F_\tau, \forall \bar{s}_\tau \in [0, 1]\}$$

$$\Xi_2 = \{\tau \in \bar{\tau} | \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(\bar{s}_\tau; (\bar{s}_\tau, \bar{s}_{-\tau}) = (\bar{s}_\tau, 0), N) < F_\tau, \bar{s}_\tau \in [0, 1]\}$$

$\Xi_1$ is the set of types that always enter a session while $\Xi_2$ is the set of types that never enter a session. Partition $\bar{\tau} = (\tau_1, \tau_2, \tau_3)$ where $\tau_1 \in \Xi_1$, $\tau_2 \in \Xi_2$, and $\tau_3 \in (\Xi_2 \cup \Xi_1)'$.

Then,

$$\tau \in \Xi_1 \Rightarrow \bar{s}_\tau = 1$$

$$\tau \in \Xi_2 \Rightarrow \bar{s}_\tau = 0$$

Update $\Xi_1$, $\Xi_2$ to incorporate these thresholds:

$$\Xi_1' = \{\tau \in \bar{\tau} | \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(\bar{s}_\tau; (\bar{s}_\tau, \bar{s}_{\tau_1}, \bar{s}_{\tau_2}, \bar{s}_{\tau_3}) = (\bar{s}_\tau, 1, 0, 0), N) > F_\tau, \forall \bar{s}_\tau \in [0,1]\}$$

$$\Xi_2' = \{\tau \in \bar{\tau} | \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(\bar{s}_\tau; (\bar{s}_\tau, \bar{s}_{\tau_1}, \bar{s}_{\tau_2}, \bar{s}_{\tau_3}) = (\bar{s}_\tau, 1, 0, 1), N) < F_\tau, \forall \bar{s}_\tau \in [0,1]\}$$

By monotonicity of profit functions we have that $\Xi_1 \subset \Xi_1'$, $\Xi_2 \subset \Xi_2'$. If the new sets differ from the originals we iterate the procedure to include new boundaries in its definition. This iteration ends because the type space is finite. Assume it ends in the first stage, without loss of generality.

Let $\tau \in (\Xi_1 \cup \Xi_2)'$. There exists $s_{1,\tau}, s_{2,\tau} \in [0,1]$ such that

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{2,\tau}; (s_{2,\tau}, 1, 0, 0), N) < F_\tau < \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{1,\tau}; (s_{1,\tau}, 1, 0, 1)), N)$$

Once more using monotonicity of $\Pi$, we get that $s_{2,\tau} > s_{1,\tau}$. Define the mapping

$$\Phi \; \Gamma \; [s_{1\tau}, s_{2\tau}]^{(\Xi_1 \cup \Xi_2)'} \to \mathbb{R}^{(\Xi_1 \cup \Xi_2)'}$$

$$\Phi_\tau(s) = \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_\tau; (s_\tau, 1, 0, s), N) - F_\tau$$

This mapping inherits its continuity from the profit functions and has the property that $\Phi_\tau(s_{1\tau}, s_{-\tau}) > 0 > \Phi_\tau(s_{2\tau}, s_{-\tau})$ for all $s_{-\tau} \in [s_{1\tau}, s_{2\tau}]^{(\Xi_1 \cup \Xi_2 \cup \{\tau\})'}$. A generalization of the intermediate value theorem, the Poincaré–Miranda theorem, ensures there exists $s^* \in [s_{1\tau}, s_{2\tau}]^{(\Xi_1 \cup \Xi_2)'}$ such that $\Phi(s^*) = 0$ which concludes the proof.

Finally, take any pure strategy equilibria. Let the set of signals for which bidders enter be denoted by $(S_\tau^*)_\tau$. Suppose this equilibrium is not in cut off strategies. There exists types $\tau$, signals $s_{1,\tau}$, a maximum signal $s_{2,\tau}$ with $s_{2,\tau} > s_{1,\tau}$ such that a bidder of type $\tau$ with signal $S = s_{2,\tau}$ enters but a bidder with $S = s_{1,\tau}$ does not.

I assert that

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{1,\tau}; (S_\tau^*)_\tau, N) = F_\tau$$

Indeed, if that is not the case we would have

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{1,\tau};(S^*_\tau)_\tau, N) < F_\tau \leq \sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s_{2,\tau};(S^*_\tau)_\tau, N) \Rightarrow s_{1,\tau} > s_{2,\tau}, \text{a contradiction.}$$

The indifference condition of players with signals $s_{1,\tau}$ and $s_{2,\tau}$ plus stochastic ordering implies that $G_{\tau,\omega}(y|s_{1,\tau}) = G_{\tau,\omega}(y|s_{2,\tau})$ a.e.

The indifference condition of players with signals $s_{1,\tau}$ and $s_{2,\tau}$ plus the convexity of probability of winning on other players' types implies that

By the same reasoning, for every $s \in [s_{1,\tau}, s_{2,\tau}]$:

$$\sum_{\omega \in \mathcal{A}_\tau} \Pi_{\tau,\omega}(s;(S^*_\tau)_\tau, N) = F_\tau$$

Thus, the types that violates cutoff strategies are indifferent from entry and it is also an equilibrium for them to enter the session. Stochastic ordering guarantees that the probability of winning is unchanged if these firms enter the session and, therefore, these two equilibria are payoff equivalent. □

## 2.8 Appendix II: Figures

### 2.8.1 Bounds on order statistics

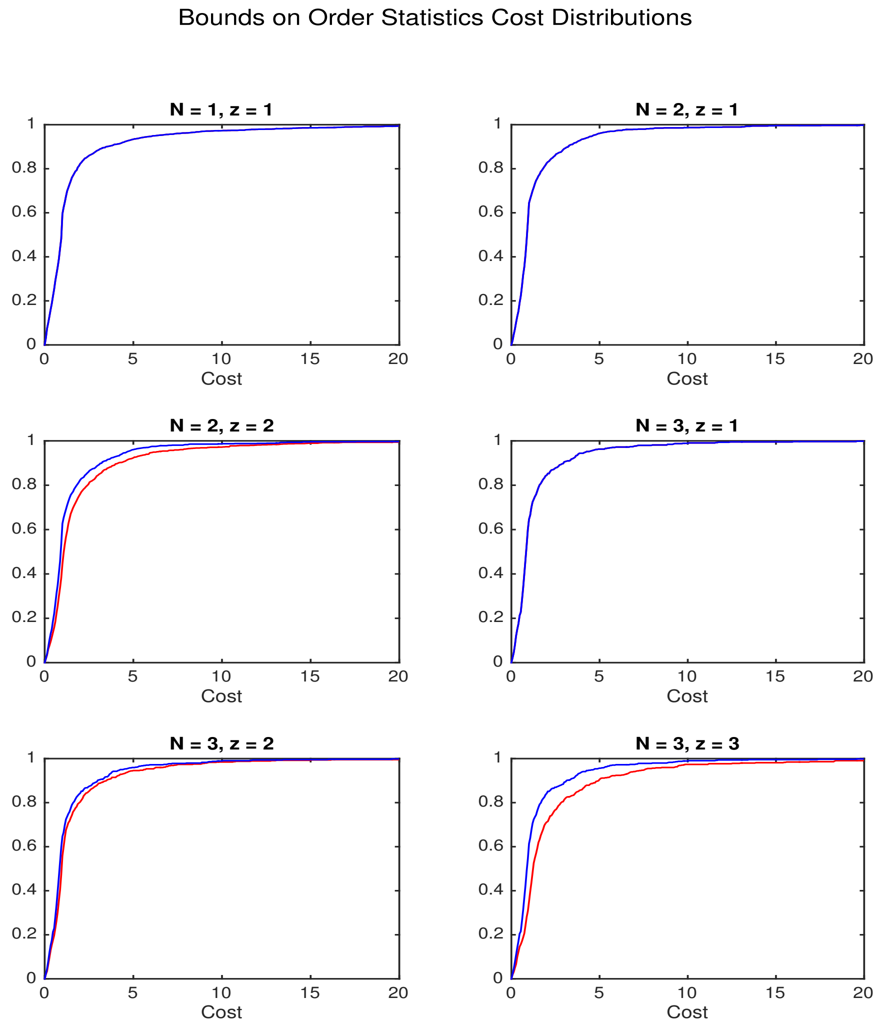



Figure 2.1: Bounds on conditional order statistics for firm type *far*. Estimated with auctions where N players entered and achieved the z order statistic with their bids. They were constructed using the residuals from section 2.4.1. Upper bounds in blue and lower bounds in red. Figure presents results for the firm type *far* and $\bar{s} = 1$. Here, the firm type is *far* if it is more than 300km away from the session.

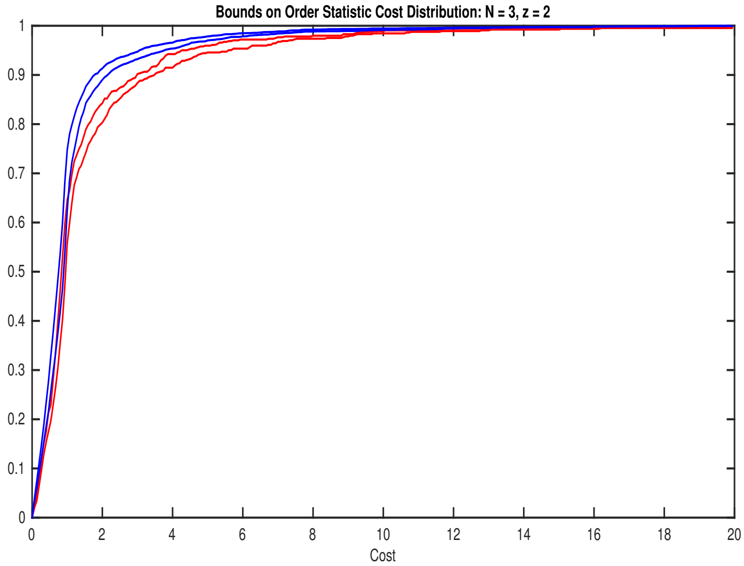



Figure 2.2: Bounds on conditional order statistics for both firm types *close* and *far*. Estimated with auctions where N=3 players entered and achieved the z=2 order statistic with their bids. They were constructed using the residuals from section 2.4.1. Firm type *close* in blue and *far* type in red. Figure presents results for $\bar{s} = 1$. Here, the firm type is *far* if it is more than 300km away from the session.

### 2.8.2 Bounds on cost distribution

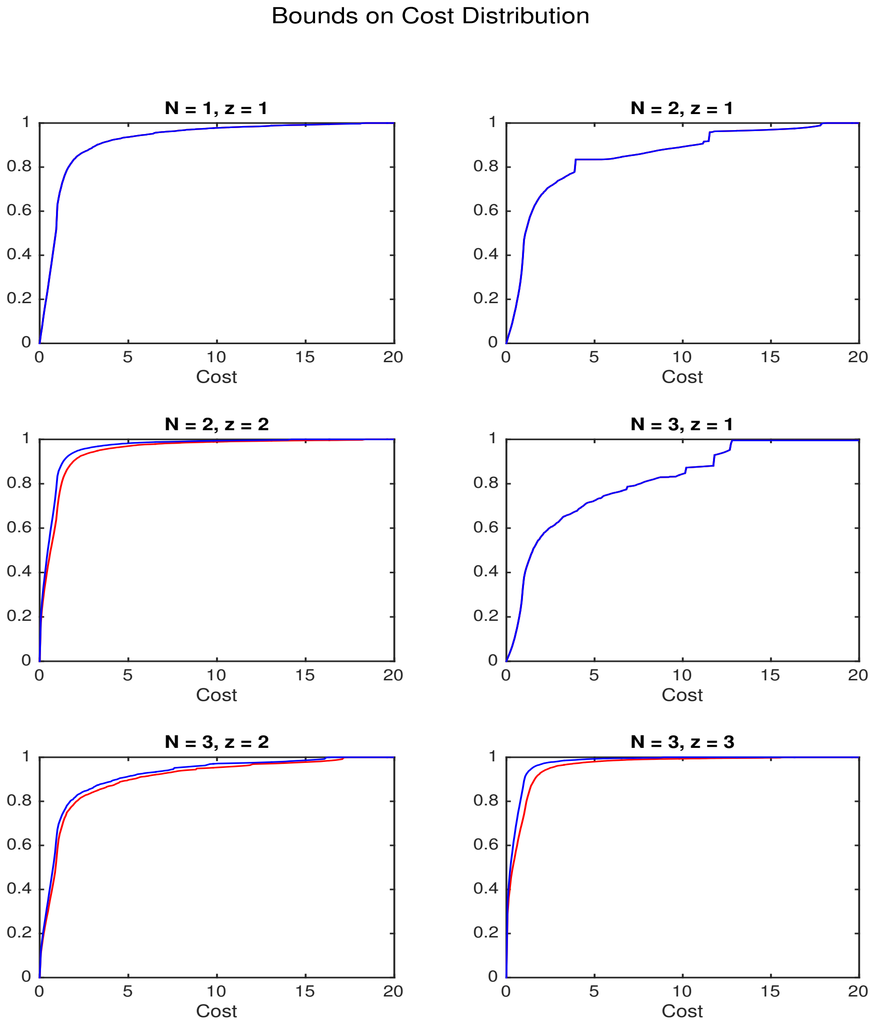



Figure 2.3: Bounds on conditional cost distribution for firm type *far*. Estimated with auctions where N players entered and achieved the z order statistic with their bids. Constructed using bounds on order statistics and the mapping between those and cost distributions (equation 2.6). Upper bounds in blue and lower bounds in red. Figure presents results for the firm type *far* and $\bar{s} = 1$. Here, firms are far if they are more than 300km away from the session.

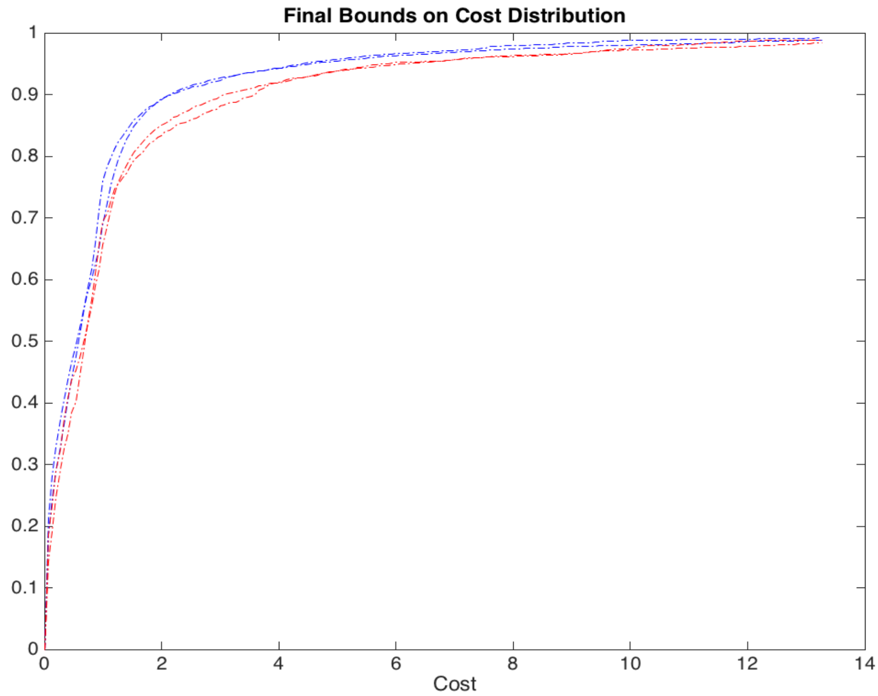



Figure 2.4: Smoothed estimators of the minimum upper bound and maximum lower bound for the conditional cost distribution. Firm type *close* in blue and *far* in red. Figure presents results estimated with auctions up to N=3, and $\bar{s} = 1$. Here, firms are far if they are more than 300km away from the session.

### 2.8.3 Shocks on entry

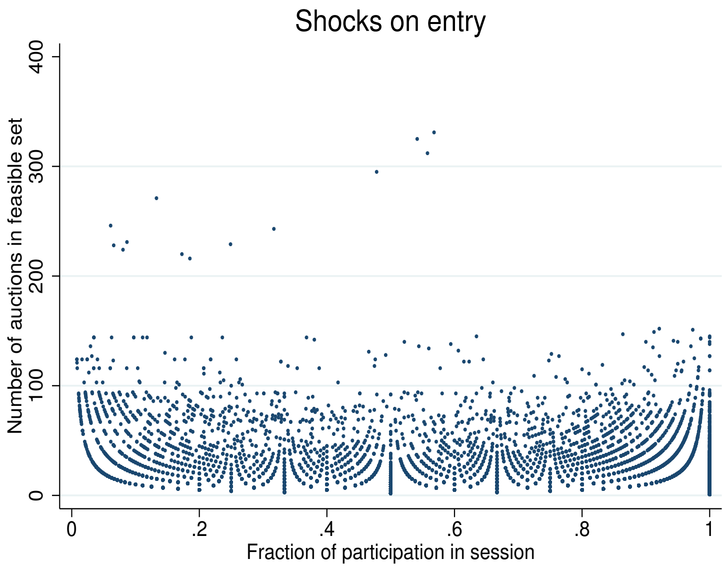



Figure 2.5: Scatter plot at the session level of the number of auctions in the firms' feasible set versus the fraction of auctions participated in. Firms feasible sets are constructed as the intersection between the firms portfolio and the menu of items auctioned at the session.

## 2.9 Appendix III: Tables

### 2.9.1 Profits

| Cutoff for firm type | Entry | Firm type | Lower Bound | Upper Bound | # Items |
|---|---|---|---|---|---|
| 100km | Yes | Far | 64.5585 | 67.1138 | 6.3280 |
| | | Close | 72.9057 | 74.6288 | 7.4703 |
| | No | Far | 47.4633 | 49.3021 | 4.9747 |
| | | Close | 38.7527 | 39.6687 | 4.5232 |
| 200km | Yes | Far | 60.5348 | 62.1760 | 6.0926 |
| | | Close | 73.7728 | 75.7374 | 7.4737 |
| | No | Far | 45.8121 | 47.062 | 4.9829 |
| | | Close | 41.7673 | 42.9192 | 5.0008 |
| 300km | Yes | Far | 58.3023 | 58.3537 | 5.8931 |
| | | Close | 74.0632 | 74.4488 | 7.3171 |
| | No | Far | 46.8202 | 46.8706 | 4.9365 |
| | | Close | 42.2676 | 42.4788 | 4.9745 |

Table 2.1: Bounds on mean session profits by entry decisions and firm type.

### 2.9.2 Entry Costs

| Cutoff for firm type | Variables | Bounds |
|---|---|---|
| 100km | $F_1$ | [11.6342,12.1261] |
| | $F_2$ | [-0.8312,-0.6052] |
| | $\alpha$ | [0.8693,0.871] |
| 200km | $F_1$ | [11.4091,11.6957] |
| | $F_2$ | [-0.5340,-0.5301] |
| | $\alpha$ | [0.8628,0.8649] |
| 300km | $F_1$ | [11.4739,11.5114] |
| | $F_2$ | [-0.3898,-0.3367] |
| | $\alpha$ | [0.8648,0.8669] |

Table 2.2: Entry cost estimates.

# CHAPTER 3

# Freemium Packages in Two-Sided Markets

## 3.1 Introduction

On many entertainment platforms, consumers interact with advertisers. Often, consumers are exposed to a freemium subscription option: they can use it for free while receiving advertising posts, or they can pay for a premium package that includes the benefit of not receiving advertising.[1]

Advertisers, on the other side of the market, choose whether to advertise on the platform. A key feature of the platform is the externality generated by the interaction of consumers with advertisers. Advertisers enjoy using the platform since it exposes their brands to a larger set of users, while these users usually dislike advertising[2]. This is an important characteristic of a two-sided market, a market in which two sets of distinct agents interact and are inherently exposed to network externalities.

Under two-sided markets, interesting scenarios might appear: the platform might choose to charge one side below its marginal cost so that it attracts more consumers to enter the market on that side and can consequently charge a higher price from the other side (assuming that it values more consumers on the other side). This scheme is observable in several markets, such as the yellow pages or subway newspapers, where consumers get to enjoy the respective goods for free. As a result, consumer welfare in a two-sided framework might differ substantially when compared to standard monopolies where a consumer gets its surplus severely extracted. Major examples of platforms that use the freemium pricing strategy include the music streaming industry (with platforms such as Spotify and Pandora), smartphone applications, and YouTube.

---

[1] premium packages might contain beneficial features in addition to no advertising, including services, and virtual/physical goods.

[2] Elena Argentesi and Lapo Filistrucchi show that this is true for newspaper platforms.

In this chapter, I study the pricing of freemium packages and its consequences to consumers in a monopolistic setting. To do so, I build a theoretical model of a platform that offer a freemium package to consumers: they choose between a free and a premium package. I characterize the degree to which the platform incorporates consumers' distaste for advertising in its pricing scheme and shed some light on the discussion of consumers' overexposure to advertising. In two extensions of the model, I provide a positive result on the identification of monopolist model fundamentals and its generalization to platform competition.

The rest of the chapter is organized as follows: Section 3.2 provides a literature review of theory and empirical applications of two-sided markets, Section 3.3 formalizes the model of a monopolistic platform and theoretical results, and Section 3.4 concludes. The extensions of the model are presented in Section 3.5 and proofs are provided in Section 3.6.

## 3.2 Literature Review

Before presenting the literature on this topic, it is important to begin with a formal definition of a two-sided market, as there is no consensus in the literature of its correct definition (Rysman (2009)). For this purpose, I follow the definition set by Rochet and Tirole (2006) and Weyl (2010):

**Definition** The market is called two-sided if it incorporates three features:

1. Multi-product firm: a platform provides a distinct service to two sides of the market, which can be explicitly charged different prices.
2. Cross network effects: users' benefit from participation depends on user participation on the other side of the market, which varies with market conditions.
3. Bilateral market power: platforms are price setters (monopolistic or oligopolistic) on both sides of the market and typically set uniform prices.

This definition is important in defining a two-sided market: the network is one-sided if firms offer a single product in the market (consumers and advertisers are indistinguishable); if there are no cross-network effects, the firm is simply a monopolist to each one of the two sides; if the platform has no market power, then it is a distributor of content. There is a rich, vast and ongoing developing literature about two-sided markets, which I summarize below.

### Theory on Two-Sided markets

This project is closely related to the theoretical developments of two-sided frameworks. Jullien & Caulliard (2003) propose a framework to analyze imperfect competition under the presence of indirect network externalities. As they demonstrate, multiple equilibria might emerge in this setting (chicken-egg problem). I show that the same type of equilibria might appear in the model considered in this chapter in the case where the two sides present complementarity with respect to their valuation of the other side in the market. Intuitively, when the agents in the two sides are willing to join the platform if, and only if, agents in the other side also join, then there are at least the two-equilibria where both or none join the platform. Multiplicity of equilibria might be an undesired result when estimating parameters of the preference of consumers or assessing counterfactuals. Jullien & Pavan (2016) show that asymmetric information on user preferences retrieves a unique equilibrium to the underlying global game. I deal with multiplicity of equilibria by solving the game in terms of quantities rather than prices.

The three most relevant contributions to the theoretical framework come from Rochet and Tirole (2003), Armstrong (2006) and Weyl (2010). These papers analyze how different platform pricing behaviors affect consumers' welfare and differ in the way they model consumers' preferences. Rochet and Tirole (2003) assume that agents have homogeneous values for the platform (membership value) and heterogeneous interactions values (network externalities), while Armstrong (2006) assumes the opposite: heterogeneous membership values and homogeneous interaction values. Weyl (2010) points out that, in general, both sources should be considered, since many types of users might be on the margin to join the platform. I build on the models offered by these authors to study the pricing of freemium packages. This allows me to relate and isolate the novelty of my results.

Finally, this is not the first attempt to study price discrimination on two-sided markets. Böhme (2012) and Liu & Serfes (2013) study second degree and third degree price discrimination, respectively. The modeling approach I follow differs substantially from Böhme. I am interested in the effects of freemium packages to consumers while he is interested in providing a descriptive analysis of the effects of price discrimination with positive prices. Moreover, in his approach, the mass of users from one side of the market is fixed, while I endogenize the participation of both sides.

## 3.3 Monopolistic Platform

There is a mass $[0,1]$ of consumers on each side of the market: side $\mathcal{A}$, represented by advertisers and side $\mathcal{B}$, represented by consumers. A single platform exerts market power on both sides. As in Weyl (2010), I assume two sources of heterogeneity for preferences of agents:

- $\theta_i^{\mathcal{I}}$ is a membership value: benefit/cost agent $i$ from side $\mathcal{I}$ get even if there is no entry on the other side of the market;
- $\eta_i^{\mathcal{I}}$ is an interaction value: benefit/cost agent $i$ from side $\mathcal{I}$ get from the interaction with the other side of the market.

A negative sign for $\eta_i^{\mathcal{I}}$ means that agent $i$ from side $\mathcal{I}$ has a lower value for the platform when more agents on the other side $\mathcal{J}$ of the market enter the platform.[3] A negative sign for $\theta_i^{\mathcal{I}}$ means that agent $i$ from side $\mathcal{I}$ bears a cost for entering the market.[4] Besides being economically intuitive, these two sources of heterogeneity will allow for a richer set of consumers to be on the margin between choosing packages and an outside option. I do not impose strong restrictions on the distribution of the heterogeneity. Rather, I assume it has a joint distribution with continuous pdf $f_I(\theta^{\mathcal{I}}, \eta^{\mathcal{I}})$, $\mathcal{I} = \mathcal{A}, \mathcal{B}$.

The platform sets prices $p^{\mathcal{A}}, p^{\mathcal{B}}$ for advertisers and consumers, respectively. Given these prices, consumers on each side choose their action simultaneously. On the one hand, advertisers can choose one of two options: (i) enter (mass $A$); (ii) outside option (mass $1-A$). On the other hand, consumers can choose one out of three options: (iii) use platform for free (mass $F$); (iv) pay price and don't receive network externalities (mass $B$); (v) outside option (mass $1-F-B$). I assume that advertisers that enter the market only interact with consumers who chose to use the platform for free.

Advertisers derive utility from joining the platform according to

$$\theta^{\mathcal{A}} + \eta^{A} \cdot F - p^{\mathcal{A}}$$

[3]This is the case if consumers dislike advertising.

[4]for instance, advertisers might have to do market research.

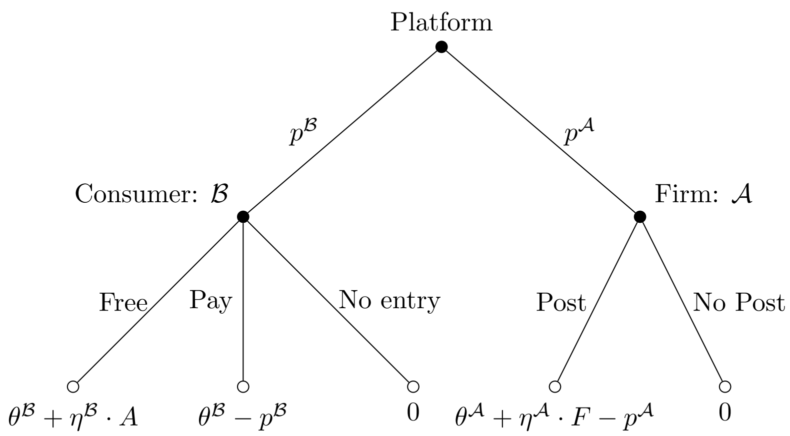
Consumers derive utility from paying for the platform according to

$$\theta^{\mathcal{B}} - p^{\mathcal{B}}$$

and from using the platform for free according to

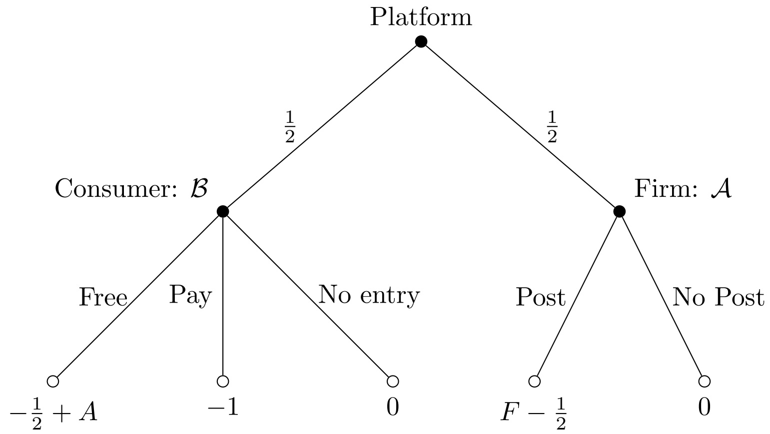
$$\theta^{\mathcal{B}} + \eta^{\mathcal{B}} \cdot A$$

The following figure represents the above-mentioned two-sided structure with corresponding payoffs:




### Example of multiplicity of equilibria

Suppose, for simplicity, that $f_{\mathcal{A}}$ and $f_{\mathcal{B}}$ are degenerate so that $\theta^{\mathcal{A}} = 0$, $\theta^{\mathcal{B}} = -1/2$, $\eta^{\mathcal{I}} = 1$, $p^{\mathcal{I}} = 1/2, \mathcal{I} \in \{\mathcal{A}, \mathcal{B}\}$. With this scheme, since both agents appreciate each other's presence in the market, no one gets to pay for the platform since the outside option is comparatively better.




Participation constrains are:

$$\begin{cases} \text{Side } \mathcal{B}: \quad A > 1/2 \\ \text{Side } \mathcal{A}: \quad F > 1/2 \end{cases}$$

This example shows that a chicken-egg equilibria might arise with all or none of the users choosing to enter the platform. To deal with the issue of multiplicity of equilibria, change the choice space of the platform from prices to the choice of quantities (share of agents on each branch of the tree). The method is detailed below.

### Side $\mathcal{A}$

Consumers from side $\mathcal{A}$ enter if they get a higher utility than the outside option, which I normalize to zero.

$$\theta^{\mathcal{A}} + \eta^{\mathcal{A}} F - p^{\mathcal{A}} \geq 0 \Leftrightarrow \theta^{\mathcal{A}} \geq p^{\mathcal{A}} - \eta^{\mathcal{A}} F$$

This means that the proportion of advertisers present in the platform is given by:

$$A = \int_{-\infty}^{\infty} \int_{p^{\mathcal{A}} - \eta F}^{\infty} f_{\mathcal{A}}(\theta, \eta) d\theta \, d\eta \tag{3.1}$$

**Side $\mathcal{B}$**

Agents choose the free package if and only if

$$\begin{cases} \theta^{\mathcal{B}} + \eta^{\mathcal{B}} A \geq \theta^{\mathcal{B}} - p^{\mathcal{B}} \Leftrightarrow \eta^{B} \geq -\frac{p^{\mathcal{B}}}{A} \\ \theta^{\mathcal{B}} + \eta^{\mathcal{B}} A \geq 0 \Leftrightarrow \theta^{\mathcal{B}} \geq -\eta^{\mathcal{B}} A \end{cases}$$

Thus, the proportion of consumers that use the platform for free is given by:

$$F = \int_{-\frac{p^{\mathcal{B}}}{A}}^{+\infty} \int_{-\eta A}^{\infty} f_{\mathcal{B}}(\theta, \eta) d\theta \, d\eta \tag{3.2}$$

Similarly, agents chooses to pay if and only if

$$\begin{cases} \theta^{\mathcal{B}} - p^{\mathcal{B}} \geq \theta^{\mathcal{B}} + \eta^{\mathcal{B}} A \Leftrightarrow \eta^{\mathcal{B}} \leq -\frac{p^{\mathcal{B}}}{A} \\ \theta^{\mathcal{B}} - p^{\mathcal{B}} \geq 0 \Leftrightarrow \theta^{\mathcal{B}} \geq p^{\mathcal{B}} \end{cases}$$

Thus,

$$B = \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}(\theta, \eta) d\theta \, d\eta \tag{3.3}$$

Using the Implicit Function Theorem I am able to write, locally, prices as a function of the share of consumers in each branch of the market. Formally, the conditions that allow for the application of the theorem are as follows:

$$\frac{\partial A}{\partial p^{\mathcal{A}}} = -\int_{-\infty}^{\infty} f_{\mathcal{A}}(p^{\mathcal{A}} - \eta F, \eta) d\eta < 0$$

$$\frac{\partial F}{\partial p^{\mathcal{B}}} = \frac{1}{A} \cdot \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta > 0$$

Fixing the share $F$, we can invert (3.1) locally as $p^{\mathcal{A}} = p^{\mathcal{A}}(A, F)$. Also, fixing share $A > 0$, we can invert equation (3.2) locally as $p^{\mathcal{B}} = p^{\mathcal{B}}(A, F)$. Thus, given a mass of consumers $\widetilde{F}$, $\widetilde{A}$ there is a unique pair of prices, profit and welfare consistent with $\widetilde{F}$, $\widetilde{A}$ users. Moreover, by equation (3.3), $B$ is uniquely pinned down. Thus, by allowing the platform to choose the shares on each side of the market we avoid the chicken-egg equilibrium problem.

### 3.3.1 Equilibrium Prices

Having rewritten the prices that each side pays in terms of the market shares, I set the problem of the platform as to choose shares $F$ and $A$ in order to maximize profits, i.e., given platform costs $c^F$, $c^{\mathcal{A}}$ and $c^{\mathcal{B}}$:

$$\max_{F,A} \Big\{ -c^F \cdot F + (p^{\mathcal{B}}(A,F) - c^{\mathcal{B}}) \cdot B(A,F) \\ + (p^{\mathcal{A}}(A,F) - c^{A}) \cdot A \Big\}$$

$$\text{s. t. } B(A,F) = \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}(\theta,\eta) d\theta\, d\eta$$

Denote by

$\epsilon^{\mathcal{I}} = -\frac{\partial I}{\partial p^{\mathcal{I}}} \cdot \frac{p^{\mathcal{I}}}{I}$ the price elasticity of demand on side $\mathcal{I}$;

$\mu^{\mathcal{I}} = \frac{p^{\mathcal{I}}}{\epsilon^{\mathcal{I}}}$ the market power platform exerts on side $\mathcal{I}$;

$\tilde{\eta}^{\mathcal{A}} = \dfrac{\int_{-\infty}^{\infty} \eta \cdot f_{\mathcal{A}}\left(p^{\mathcal{A}} - \eta F, \eta\right) d\eta}{\int_{-\infty}^{\infty} f_{\mathcal{A}}\left(p^{\mathcal{A}} - \eta F, \eta\right) d\eta}$ the average interaction value of marginal users on side $\mathcal{A}$;

$\tilde{s} = -\frac{\partial F}{\partial B}$ the marginal change in $F$ due to a marginal change in $B$;

$\epsilon_{F,B} = -\frac{\partial F}{\partial B} \cdot \frac{B}{F} = -\frac{e^F}{e^B}$ the relative elasticity of demand between $F$ and $B$.

The following proposition characterizes the equilibrium prices for a profit-maximizing platform.

**Proposition 3.3.1.** *The interior profit-maximizing pricing under freemium packages is characterized by equations 3.1, 3.2 and the following two conditions:*

$$\frac{p^{\mathcal{A}} - c^{A}}{p^{\mathcal{A}}} = \frac{1}{\epsilon^{\mathcal{A}}} - \epsilon_{F,B} \cdot \frac{p^{\mathcal{B}} \cdot F}{p^{\mathcal{A}} \cdot A} \tag{3.4}$$

$$\frac{p^{\mathcal{B}} - c^{B}}{p^{\mathcal{B}}} = \frac{\epsilon_{F,B}}{\epsilon^{\mathcal{B}}} + \frac{\tilde{s}}{p^{\mathcal{B}}} \cdot [\tilde{\eta}^{\mathcal{A}} A - c^F] \tag{3.5}$$

*or, equivalently, by:*

$$p^{\mathcal{A}} = c^{\mathcal{A}} + \mu^{\mathcal{A}} - \epsilon_{F,B} \cdot \frac{p^{\mathcal{B}} \cdot F}{A} \tag{3.6}$$

$$p^{\mathcal{B}} = c^{\mathcal{B}} - \tilde{s} \cdot c^{F} + \mu^{\mathcal{B}} \cdot \epsilon_{F,B} + \tilde{s} \cdot \tilde{\eta}^{\mathcal{A}} A \tag{3.7}$$

*Proof.* See appendix 3.8.1. □

Equations (3.4) and (3.5) show that the optimal pricing follows a variation of Lerner's formula, while equations (3.6) and (3.7) are more suitable for interpretation.

$$p^{\mathcal{A}} = \underbrace{c^{\mathcal{A}}}_{\text{marginal cost}} + \overbrace{\mu^{\mathcal{A}}}^{\text{market power}} - \underbrace{\epsilon_{F,B} \cdot \frac{p^{\mathcal{B}} \cdot F}{A}}_{\text{forgone revenue from side } \mathcal{B}}$$

The first two terms that appear in equation (3.6) are standard. The degree to which the platform charges advertisers above its marginal cost depends on the market power of the platform, which is the price over elasticity of demand. Intuitively, the price will be higher when consumers have smaller demand elasticity. i.e., when their demand shifts less when facing a higher price. The last term is novel. The platform internalizes that a share of agents use the platform for free and, as a consequence, it forgoes some revenue from the consumer side. This increases the value for advertisers and allows the platform to charge them to compensate for the loss in consumer revenue.

$$p^{\mathcal{B}} = \underbrace{c^{\mathcal{B}} - \tilde{s} \cdot c^{F}}_{\text{marginal cost}} + \overbrace{\mu^{\mathcal{B}} \cdot \epsilon_{F,B}}^{\text{elastic market power}} + \underbrace{\tilde{s} \cdot \tilde{\eta}^{\mathcal{A}} A}_{\text{internalized externality from } \mathcal{A}}$$

A marginal increase in $B$ has direct effects on $F$. To start, the higher the elasticity of the share of agents using the platform for free with respect to the share of agents that pay for it, the higher is the market power that the platform imposes on side $B$. Moreover, increasing $B$ marginally implies a marginal decrease of $\tilde{s}$ on $F$, which means that the platform saves $\tilde{s} \cdot c^{F}$. Finally, the platform internalizes part of the externality generated by the interaction of advertisers and consumers. More specifically, it internalizes the effect of the average interaction value of marginal users on side $\mathcal{A}$. It employs this strategy instead of weighting the average user because, as a profit maximizer, it cares about the agent who is at the margin of joining each plan. Thus, if advertisers on the margin like interacting with consumers, the platform will then increase its price on consumers so that it shifts some of them to the free package. There is a limit to this effect since some agents also shift to the outside option. This rationalizes why the term is weighted by the marginal effect of B on F.

### 3.3.2 Comparison with Social Planner

The platform creates value for the agents that use it. In the presence of externalities, a Social Planner can set a tax to correct for inefficient outcomes. This is also known as Pigouvian or Efficient pricing. It does so by determining the tax such that the platform's marginal value is equal to its marginal cost. Intuitively, the social planner will internalize the average interaction value of consumers, as it cannot observe each agent's interaction value. Rather, the platform only knows the distribution of heterogeneity.

To determine the total value each side of the market has, I integrate out the benefit that agents get when choosing an action and subtract the costs faced by the platform. Formally,

**Side $\mathcal{A}$**

$$V^A = \int_{-\infty}^{\infty} \int_{p^{\mathcal{A}} - \eta F}^{\infty} \{\theta + \eta F\} f_{\mathcal{A}}(\theta, \eta) d\theta \, d\eta \tag{3.8}$$

**Side $\mathcal{B}$**

$$V^F = \int_{-\frac{p^{\mathcal{B}}}{A}}^{\infty} \int_{-\eta A}^{\infty} \{\theta + \eta A\} f_{\mathcal{B}}(\theta, \eta) d\theta \, d\eta \tag{3.9}$$

$$V^B = \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} \int_{p^{\mathcal{B}}}^{\infty} \{\theta\} f_{\mathcal{B}}(\theta, \eta) d\theta \, d\eta \tag{3.10}$$

### Efficient pricing

A benevolent social planner would choose shares $F$ and $A$ in order to maximize the total value of the platform, i.e.

$$\max_{F,A}\left\{V^A + V^F + V^B - c^A A - c^F F - c^B B\right\}$$

$$\text{s. t. } B(A,F) = \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}(\theta,\eta)d\theta\, d\eta$$

Denote by

$\bar{\eta}^A = \frac{\int_{-\infty}^{\infty}\int_{p^{\mathcal{A}}-\eta F}^{\infty} \eta f_{\mathcal{A}}(\theta,\eta)d\theta\, d\eta}{\int_{-\infty}^{\infty}\int_{p^{\mathcal{A}}-\eta F}^{\infty} f_{\mathcal{A}}(\theta,\eta)d\theta\, d\eta}$ the average interaction value of average users on side $\mathcal{A}$

The following result characterizes the price under this scenario.

**Proposition 3.3.2.** *The interior efficient pricing under freemium packages is characterized by equations (3.1), (3.2) and the following two conditions:*

$$p^{\mathcal{A}} = c^{\mathcal{A}} \tag{3.11}$$

$$p^{\mathcal{B}} = c^{\mathcal{B}} - \tilde{s}\cdot c^F + \tilde{s}\cdot \bar{\eta}^{\mathcal{A}} A \tag{3.12}$$

*Proof.* See appendix 3.8.2. □

Remarkably, despite the complicated interaction between the two sides, (3.11) shows that the efficient pricing consists of charging advertisers the platform's marginal cost. This is a consequence of

$$\frac{\partial V^F}{\partial A} + \frac{\partial V^B}{\partial A} = 0$$

Intuitively, this condition tells us that, given a marginal change in the share of advertisers, whatever value is created to agents that chose the free package is destroyed for agents that chose to pay for the platform. Moreover, the existence of the two packages in side $B$ allows the social planner to internalize the externality generated by the interaction of the two-sides without distorting the price advertisers pay.

On the other hand, the platform internalizes the effect of the average interaction value on average users on side $\mathcal{A}$, given by the last term in (3.12). This means that consumers will be taxed if the average advertiser that entered the market enjoys the presence of consumers, thus shifting a mass of them to the free package. This is a key difference between the social planner platform and the profit maximizing platform: the first is concerned with the *average* user, while the second is concerned with the *marginal* user. The difference between those interaction values is the *Spence distortion*. This effect represents the failure of the profit maximizing platform to internalize part of the externality generated by the interaction of the two sides. For example, music streaming platforms, such as Spotify, might be more concerned with advertisers on the margin than the average advertiser that posts on it. Assuming that the advertisers on the margin have a higher interaction value than the average advertiser, the price of the paid package on side $B$ would be distorted up, and agents on the free package would be overly exposed to advertising.

Rewriting equation (3.7), we have

$$p^{\mathcal{B}}_{profitmax} = p^{\mathcal{B}}_{efficient} + \underbrace{\mu^{\mathcal{B}} \cdot \epsilon_{F,B}}_{\text{elastic market power}} + \underbrace{\tilde{s} \cdot (\tilde{\eta}^{\mathcal{A}} - \bar{\eta}^{\mathcal{A}})A}_{\text{Spence distortion}}$$

This is similar to Weyl (2010), who shows that the profit maximization price is the socially optimal price distorted by market power and the Spence distortion. The distinction is that we consider the effects that pricing side $\mathcal{B}$ have on the choices available for it, represented by the elasticity $\epsilon_{F,B}$ and $\tilde{s}$ in (3.12). On the other hand, the price faced by advertisers is not distorted by the presence of network effects at all. Comparing (3.11) and (3.6) we see that the profit maximization price incorporates the platform market power and the fact that it foregoes some revenue by giving agents the option to use it for free.

$$p^{\mathcal{A}}_{profitmax} = p^{\mathcal{A}}_{efficient} + \underbrace{\mu^{\mathcal{A}}}_{\text{market power}} + \underbrace{\epsilon_{F,B} \cdot \frac{p^{\mathcal{B}} \cdot F}{A}}_{\text{Forgone revenue from side } \mathcal{B}}$$

Having derived that the Pigouvian prices advertisers pay equal the platform's marginal cost, suggests that it is efficient for the platform to flood consumers that use it for free with advertising, increasing the relative value of the paid version and allowing it to charge a higher price from consumers that pays for it. In the face of these results, one recommendation is to allow consumers to customize which ads are displayed for them. This would allow the platform to price discriminate on the advertisers' side and generate enough value to the platform that it possibly would not overexpose customers to advertisement.

## 3.4 Conclusion

This chapter has focused on freemium packages, a typical source of price discrimination used by platforms. I built a theoretical model of a monopolistic platform that offers consumers a choice between a free package, in which consumers are exposed to advertising, and a premium package, in which they pay not to be exposed to advertisements. This allows me to characterize the interior price for the profit maximizing and social planner platforms.

The theoretical results provide insight on how the platform internalizes externalities generated by the interaction of the two sides, as well as how it trades off revenue from the two sides of the market and the forgone revenue from free users. By giving the possibility for agents to use the platform for free, I show the direct link with how this spills over to the price advertisers and consumers pay under profit maximization. Moreover, as long as the interior prices are optimal, I show that it is efficient to charge advertisers the platform's marginal cost, suggesting that if the objective of the platform is to maximize value, consumers might be overexposed with advertising and the premium plan might feature high prices.

In the three appendices that follow, Section 3.5, 3.6, and 3.7, I provide additional results that depart from the model presented so far. First, motivated by data applications, I provide a positive result on the semi-parametric identification of value distributions of the monopolistic model. Second, I show that the model reconciles that some platforms choose to only offer a single free plan to consumers while others choose to offer a single paid version. In the context of freemium packages, these are both corner solutions of the platform's problem. Finally, I extend propositions 3.3.1 and 3.3.2 to the context of platform competition taking into account that most platforms in the market encounter themselves under oligopoly rather than monopoly.

Throughout the chapter, I assumed that the platform produces it own content. Even though this might be true for part of the television or movie-streaming markets, the assumption is not reasonable for music-streaming platforms who pays artists to stream their songs. A natural extension to the model presented here would be to incorporate a market for content, where the platform would bargain with a third side of the market (musicians), to include their products on the platform. This could be interesting because it would reconcile the fact that some artists refuse to enter agreements with platforms - either because they would be underpaid or their copyright content would be given away for free - with the double margin problem that arises from the vertical structure of platforms charging consumers and paying musicians.

## 3.5 Appendix I: Identification

The pricing theory on two-sided markets is closely linked to agents' sources of heterogeneity, especially to the network effects that arise from the interaction from the side of consumers and the side of advertisers. In the theoretical model developed in this chapter, consumers sort themselves into a free package, in which they are exposed to network effects - these arise from interaction with advertisers, and a premium package, in which they pay to avoid these effects.

In this subsection, I investigate the semi-parametric identification of network effects. The three propositions, presented below, show that these can be identified from data under a reasonable set of assumptions. The positive identification result is important for policy implications and platform design. Under the presence of externalities, a measure of welfare that only considers how much consumers pay for the subscription is not suitable. For example, it is not clear what the consumer surplus will be when there is a set of agents that use the platform for free and are exposed to advertising. From a revealed preference argument, observing this choice means that, for them, this action is better than paying for the platform or the outside option. Thus, knowing the distribution of the distaste for advertising helps the researcher, policy maker and owner of the platform to better assess welfare effects within this context.

### Empirical Literature on Two-Sided Markets

Following the developments of the theory on two-sided markets, several authors studied its empirical applications. I add to the empirical literature on two-sided markets by providing identification results of a general framework that can be taken to the data.

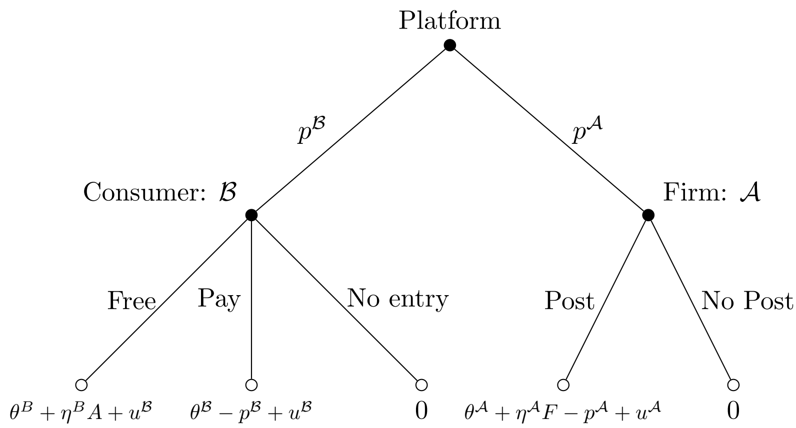
It is common to see applications using the newspaper industry, due to data availability. Argentesi & Filistrucchi (2007), for instance, estimate market power in the Italian newspaper industry to assess whether observed prices are consistent with collusion or profit-maximizing behavior. On the other hand, Chandra & Collard-Wexler (2009) study the effects of mergers in the Canadian newspaper industry. The authors show that greater concentration of the market did not lead to higher prices for either newspaper subscribers or advertisers. Angelucci & Cage & Nijs (Working Paper) develop a model to study the key determinants of second-degree price discrimination in the newspaper industry, where a company must attract both readers and advertisers. I differ from them to the extent that readers on one side of the market have heterogeneous values from reading the newspaper but do not interact with advertisers. Only advertisers have a heterogeneous taste for readers. They calibrate the model with data from local French newspapers between 1960 and 1974 and provide evidence for increased price discrimination as a result of decreased advertisement revenues.

The Yellow Pages market is a key example of two-sided markets as consumers benefit from not paying to use the platform, but are over-exposed to advertising. Busse & Rysman (2005) examine this market that emphasizes the importance of network effects showing how companies respond to increased competition by offering discount rates for advertising. Note that consumers that especially dislike ads would be willing to pay to get rid of them, but do not have this option. In my framework, agents are allowed to either use the platform for free and receive network effects, or pay for it's use and not get exposed to them.

Jeziorski (2011) analyzes mergers in the radio industry, reinforcing the point that additional market power gained from mergers might actually benefit a set of consumers (listeners). The author quantifies and disentangles the welfare effects into changes in the variety of products available to the consumer and changes in the supplied quantity of advertising.

### Game

The following figure represents a modification from the game of Section 3.3 to incorporate unobservable variables to the econometrician. Note that the outside option is kept normalized at zero utility level.



where $u = (u^{\mathcal{A}}, u^{\mathcal{B}})$ is a random vector of latent variables with distribution $F_u$ and unobservable for the econometrician.

### Consumers

The following assumptions provide sufficient conditions to identify the distribution of distaste for advertising ($F_\eta$) and index $\beta$. Let $X$ represent a vector of consumer characteristics.

**Assumption 1.** $\{B, A, p^{\mathcal{B}}, X\}$ *is observed and* $0, e_1, ..., e_K \in support\{X\}$.

**Assumption 2.** $\theta^B = X \cdot \beta$.

**Assumption 3.** $(X, \eta, u^{\mathcal{B}})$ *are mutually independent.*

**Assumption 4.** $(\eta, u^{\mathcal{B}}) \perp X$.

**Assumption 5.** $||\beta||= 1$, $dim(\beta) = K > 1$.

**Assumption 6.** $F_{u|X}$ *is strictly increasing.*

**Assumption 7.** $x^1$ *has full support on the real line and* $\beta_1 > 0$.

The first assumption requires that we observe prices, characteristics of consumers and shares of agents that choose to pay for the platform, as well as the share of advertisers that decide to post on it. The second assumption implies that I can project the value that each agent has for the platform on their observable characteristics. Even though this assumption transforms the model into semi-parametric, it is still fairly general as I do not restrict the number of observables that can be included.

Assumption (3) requires that the shocks that a consumer receives are mutually independent from his/her distaste of advertising and his/her observables. A shock might increase or decrease the overall utility of an agent, but it does not affect how much he/she enjoys the presence of side A on the platform.

Assumptions (4) and (5) are necessary for the identification of $\beta$ due to the single index nature of the problem. I follow Ichimura (1988) and Powell et al. (1986) to first identify $\beta$ and then proceed to the identification of distributions $F_\eta$ and $F_{u^{\mathcal{B}}}$.

**Proposition 3.5.1** (Identification at infinity)**.** *Suppose assumptions 1-6 hold. Then, $\beta$, $F_{\eta^{\mathcal{B}}}$ and $F_{u^{\mathcal{B}}}$ are identified.*

*Proof.* See appendix 3.8.4. □

At first glance, this is a positive result with a strong assumption on the support of the observables. However, note that I did not require that the share of agents that use the platform for free is known. Thus, Proposition 4.1 provides an identification result with partially observable data. Nevertheless, substituting assumptions (1) and (6) with (7), I show a stronger identification result that does not rely on a large support of observable characteristics of consumers.

**Assumption 8.** $\{B, F, A, p^{\mathcal{B}}, X\}$ *is observed and* $0, e_1, ..., e_K \in$ *support*$\{X\}$.

**Proposition 3.5.2.** *Suppose assumptions 2-5, 7 hold. Then, $\beta$, $F_\eta$ and $F_{u^{\mathcal{B}}}$ are identified.*

*Proof.* See appendix 3.8.5. □

The proof of the result shows that the identification of $F_\eta$ and $F_{u^{\mathcal{B}}}$ is constructive, i.e., there is a closed-form solution for the distributions of interest. Denote by $F(x_i, p^{\mathcal{B}}, A)$ and $B(x_i, p^{\mathcal{B}}, A)$ the shares of agents that choose the free and paid package, respectively, conditional on $X = x_i$ and facing prices $p^{\mathcal{B}}$ and $A$ advertisers. Then,

$$F_{u^{\mathcal{B}}}(p^{\mathcal{B}} - x_i \cdot \beta) = 1 - B(x_i, p^{\mathcal{B}}, A) \cdot \exp\left\{ \int_{-\infty}^{p^{\mathcal{B}}} \frac{\frac{\partial F(x_i,s,A)}{\partial s}}{B(x_i, s, A)} ds \right\} \tag{3.13}$$

$$F_\eta(\eta) = 1 - \exp\left\{ - \int_{-\infty}^{\eta A} \frac{\frac{\partial F(x_i,s,A)}{\partial s}}{B(x_i, s, A)} ds \right\} \tag{3.14}$$

### Discussion on Estimation

The right hand side of (3.13) and (3.14) are observed feasible to be estimated if the econometrician observes $\{p_j^{\mathcal{B}}, F_j, B_j, A_j\}_{j=1}^N$ for a fixed $x_i$. Thus, given an estimate of $\beta$, $F_{u^{\mathcal{B}}}(\cdot)$ and $F_\eta(\cdot)$ can also be estimated. If the econometrician knows $F_{u^{\mathcal{B}}}$ then she can infer $\beta$. Define $\mathbb{X} = [x_i]$ the matrix containing the data of observable characteristics. Then,

$$\beta = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}' \cdot \underbrace{\left[ p^{\mathcal{B}} - F_{u^{\mathcal{B}}}^{-1}\left( 1 - B(X, p^{\mathcal{B}}, A) \cdot \exp\left\{ \int_{-\infty}^{p^{\mathcal{B}}} \frac{\frac{\partial F(X,s,A)}{\partial s}}{B(X,s,A)} ds \right\} \right) \right]}_{\text{Kx1}} \tag{3.15}$$

Even though this formula is very suitable, in practice it is unlikely that one knows $F_{u^{\mathcal{B}}}$. Thus, it is necessary for us to develop a method to estimate $\beta$ and $F_{u^{\mathcal{B}}}$ simultaneously.

## Preliminary idea

By Assumption 1, $X = 0$ is observed for some agent. This assumption means that he/she only gets utility through a positive interaction value or through a positive shock. Otherwise, the consumer would not enter the platform. By (3.13):

$$F_{u^{\mathcal{B}}}(p^{\mathcal{B}}) = 1 - B(0, p^{\mathcal{B}}, A) \cdot \exp\left\{ \int_{-\infty}^{p^{\mathcal{B}}} \frac{\frac{\partial F(0,s,A)}{\partial s}}{B(0,s,A)} ds \right\} \tag{3.16}$$

It is possible to recover $F_u^{\mathcal{B}}$ through variation of prices. One could simulate the right hand-side of the (3.16) for several prices, use an interpolation method as the spline, mentioned above, or use a Kernel estimation for the cumulative distribution. Another approach would be to adapt Ichimura's single index estimator for this context and estimate $\beta$ and $F_{u^{\mathcal{B}}}$ simultaneously. Once an estimate $\hat{F}_{u^{\mathcal{B}}}$ is obtained, $\beta$ can be estimated by:

$$\hat{\beta} = (\mathbb{X}'\mathbb{X})^{-1}\mathbb{X}' \cdot \left[ p^{\mathcal{B}} - \hat{F}_{u^{\mathcal{B}}}^{-1}\left( 1 - B(X, p^{\mathcal{B}}, A) \cdot \exp\left\{ \int_{-\infty}^{p^{\mathcal{B}}} \frac{\frac{\partial F(X,s,A)}{\partial s}}{B(X,s,A)} ds \right\} \right) \right] \tag{3.17}$$

### Advertisers

A similar set of assumptions ensures sufficient conditions needed to identify the distribution of the network effects ($F_\eta$) and index ($\xi$).

**Assumption 9.** $\{F, A, p^{\mathcal{A}}, Y\}$ *is observed and* $0, e_1, ..., e_K \in support\{Y\}$.

**Assumption 10.** $F_{\eta^B} = F_{\eta^A} = F_\eta$.

**Assumption 11.** $\theta^A = -Y \cdot \xi$.

**Assumption 12.** $(Y, \eta, u^{\mathcal{A}})$ *are mutually*

**Assumption 13.** $||\xi||= 1$.

**Assumption 14.** $F_{u^{\mathcal{A}}+\eta \cdot F}$ *is strictly increasing.*

Assumption 9 has an important role. It means that consumers and advertisers draw from the same distribution no matter how much they like/dislike each other's presence on the platform. If I drop this assumption, then I can only identify the distribution of $u^{\mathcal{A}} + \eta^A \cdot F$. This is due to the fact that on side $\mathcal{A}$ there is only one equation relating how many advertisers chose to post on the platform with the distributions of interest. An alternative approach is to consider Assumption 9':

**Assumption 9'.** $F_{u^{\mathcal{B}}} = F_{u^{\mathcal{A}}} = F_u$.

**Proposition 3.5.3.** *Suppose assumptions 2-5 and 7-13 hold. Then,* $\xi$, $F_\eta$ *and* $F_{u^{\mathcal{A}}}$ *are identified.*

*Proof.* See appendix 3.8.6. □

The weaker identification result on the advertisers' side comes from the weaker structure of the model on this side. Nevertheless, there is no loss in terms of measuring their welfare, as the profit of the advertisers is a better way to measure their surplus.

## 3.6 Appendix II: Simulation

Having characterized the interior solution for the profit maximizing platform in proposition 3.3.1, I run a simulation to show that the proposed model can reconcile different platform pricing behaviors.

The motivation for this exercise is due to the empirical fact of some platforms offering a single free version or a single paid version of their product, which are corner solutions to the platform problem considered here. For instance, the platform making revenue only through advertisers is associated with all consumers choosing the free package, i.e., $F = 1$.

I consider a simple parametric specification for the distribution of heterogeneity of each side:

$$f_{\mathcal{B}}(\theta, \eta) = \mathbb{1}_{[0,1]X[-1,0]}$$

$$f_{\mathcal{A}}(\theta, \eta) = \frac{1}{ad}\mathbb{1}_{[-a,0]X[0,d]}$$

Side $\mathcal{B}$ has a fixed joint Uniform distribution on $[0, 1]X[-1, 0]$ and side $\mathcal{A}$ has a joint Uniform on $[-a, 0]X[0, d]$. The objective of the exercise is to analyze the profit obtained by the platform by varying parameters $(a, d) \in [0, 4]X[0, 8]$. The maximum values for $a$ and $d$ are chosen in order to allow for a simple comparison of the two-sided with a monopolist framework, as discussed below. Importantly, notice that this specification assumes that all consumers (side $\mathcal{B}$) dislike advertising and have a positive membership value while all advertisers enjoy the presence of consumers on the platform and have a fixed cost to post an advertisement.

The costs are parametrized as follows:

$$c^F = c^{\mathcal{B}} = 0.02, c^{\mathcal{A}} = 0.03$$

I also chose cost values so that the comparison between frameworks is simple and intuitive. Other values were also tested and produced similar results to the ones shown below. Figure 3.1 plots the profit obtained from an interior solution to the two-sided market versus standard monopoly for the parameter range. Notice that for low values of $a$ monopoly is better while on the other region two-sided generates higher profits.

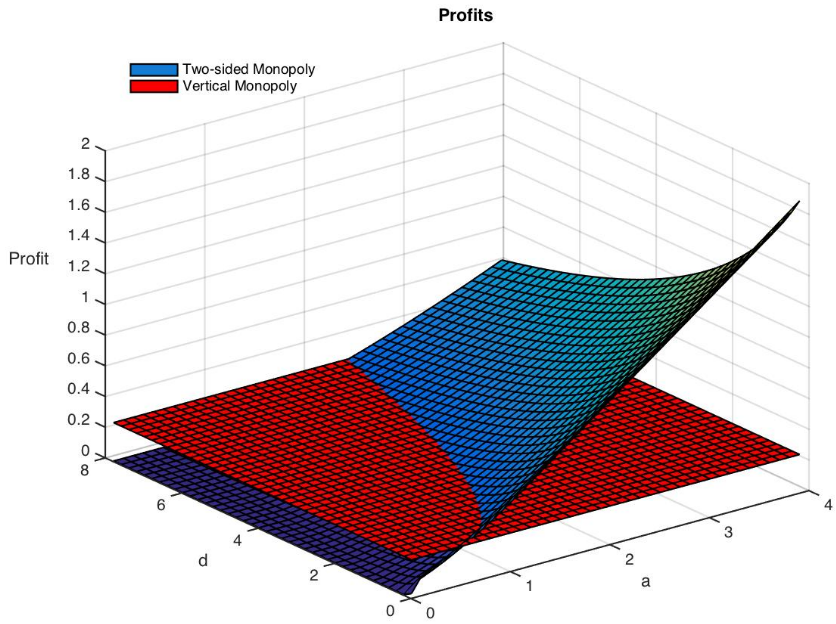



Figure 3.1: Platform Profits: Two sided vs One sided

Next, I show the comparison between prices that consumers pay under standard monopoly and two-sided monopoly. Figure 3.2 shows that consumers face a higher price under two-sided when $a$ is large and $d$ is small. In this region a high proportion of advertisers face a high cost for a post and small interaction benefit. As a consequence, the platform charges a high price for consumers with the objective to incentivize a fraction of consumers to migrate to the free package - so that they increase the revenue from advertisers, and also make a high revenue for agents that remain in the paid package. Figure 3.1 shows that this is the region where the platform makes the highest profit.

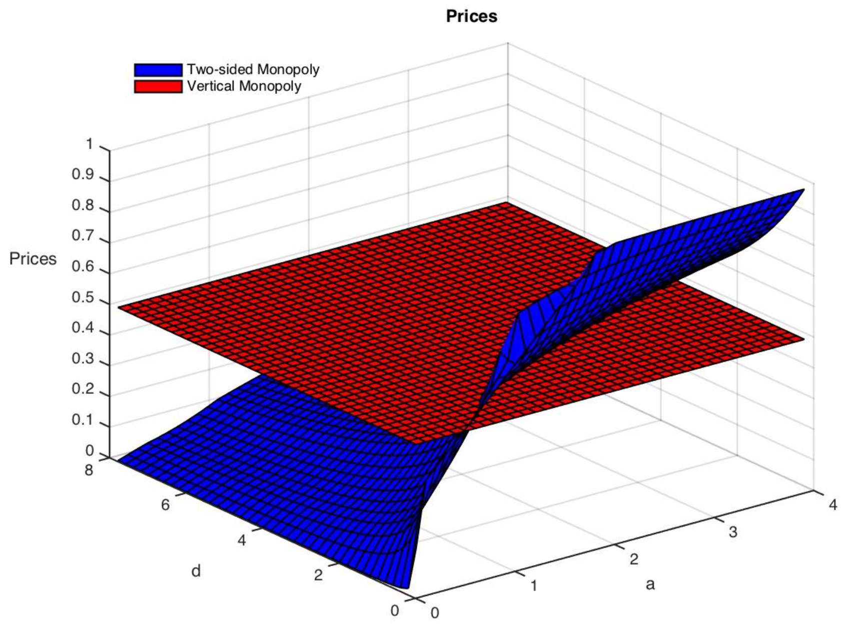



Figure 3.2: Consumer Prices: Two sided vs One sided

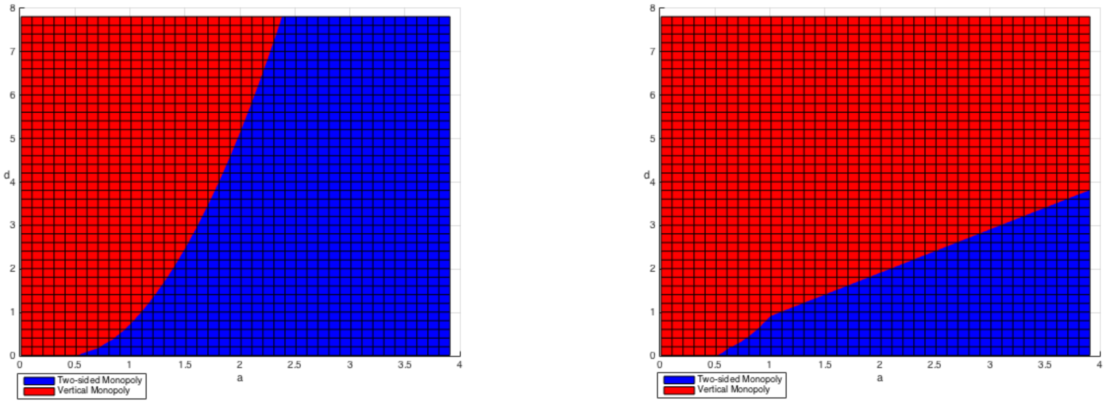
The view from above Figure 3.1 and Figure 3.2 put side by side in Figure 3.3, shows whether consumers face a higher price under two-sided or monopoly, for each choice of parameter, and the framework that allows the platform to get higher profits. Overlaying the two figures, as in Figure 3.4, four regions emerge:

I. Monopoly with higher consumer price;

II. Two-sided with lower consumer price;

III. Two-sided with higher consumer price;

IV. Monopoly with lower consumer price.




(a) Platform Profits (b) Consumer Prices

Figure 3.3: Figure 3.1 and Figure 3.2: views from above

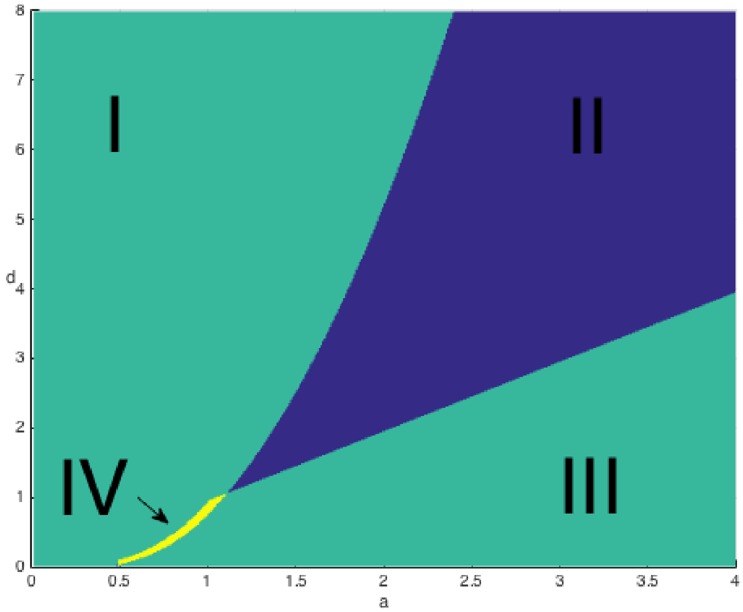



Figure 3.4: Figure 3.1 and Figure 3.2 overlaid

These regions point out that the model reconciles the fact that platforms might choose different pricing strategies depending on user preferences. For instance, if advertisers have a distribution of heterogeneity such that they belong to region II, then the two-sided market pricing strategy is best. On the other hand, if they belong to region I, then monopoly pricing (corner solution of the two-sided model) is best and the platform should forego advertising. Moreover, note that region II is beneficial to consumers. This is because the market allows them to choose between a free and paid plan priced lower than the standard monopoly.

## 3.7 Appendix III: Platform Competition

So far, I have assumed the existence of a single platform in the market. This allowed us to isolate and understand the economic forces that are incorporated in the pricing decision of the company. From an empirical point of view, it would be productive to incorporate competition effects on this model, as most platforms encounter themselves in an oligopoly market. With this point in mind, I now develop a model of platform competition with two-sided markets, focusing on the freemium packages.

The first assumption that is necessary to be addressed, is what decisions agents are now allowed to make. In practice, both sides could join any of the $J$ platforms. For example, a consumer could have an account on both Spotify and Pandora. For simplicity, I assume that advertisers can post on several platforms while agents can choose at most one platform. This characterizes a *bottleneck* model, where one side of the market sets multi-homes and the other sets a single-home.

Next, I assume that agents draw their membership and interaction values from a distribution with continuous pdf $f_{\mathcal{I}}(\theta_1, ..., \theta_J, \eta_1, ..., \eta_J)$. This means that their tastes are heterogeneous across platforms.

### Side $\mathcal{A}$

Advertisers derive utility from joining platform $j$ according to

$$\theta^{\mathcal{A}} + \eta_j^{\mathcal{A}} F_j - p_j^{\mathcal{A}}$$

It is optimal to post on platform $j$ as long as the derived utility is higher than the outside option, normalized to zero here. Denote $\eta = (\eta_1, ..., \eta_J)$,

$$A_j = \int_{-\infty}^{\infty} \int_{p_j^{\mathcal{A}} - \eta_j F_j}^{\infty} f_{\mathcal{A}}(\theta, \eta) d\theta \, d\eta \tag{3.18}$$

### Side $\mathcal{B}$

Consumers derive utility from paying for platform $j$ according to

$$\theta_j^{\mathcal{B}} - p_j^{\mathcal{B}}$$

while it derives utility from using platform $j$ for free according to

$$\theta_j^{\mathcal{B}} + \eta^{\mathcal{B}} A_j$$

An agent chooses the free package of $j$ platform if and only if: (i) it is better than the outside option; (ii) it is better than any other free package; (iii) it is better than any paid package. That is:

i $\theta_j^{\mathcal{B}} + \eta^{\mathcal{B}} A_j \geq 0$

ii $\theta_j^{\mathcal{B}} + \eta^{\mathcal{B}} A_j \geq \theta_k^{\mathcal{B}} + \eta^{\mathcal{B}} A_k, \forall k \neq j$

iii $\theta_j^{\mathcal{B}} + \eta^{\mathcal{B}} A_j \geq \theta_k^{\mathcal{B}} - p_k^{\mathcal{B}}, \forall k$

Define $V_j$ being the region on $\mathbb{R}^{J+1}$ defined by these constrains and $\theta = (\theta_1, ...\theta_J)$. Then, the proportion of consumers that use the platform for free is given by:

$$F_j = \int_{V_j} f_{\mathcal{B}}(\theta, \eta) d\theta \, d\eta \tag{3.19}$$

Likewise, an agent chooses the paid package of $j$ platform if and only if: (iv) it is better that the outside option; (v) it is better than any other paid package; (vi) it is better than any free package. That is:

iv $\theta_j^{\mathcal{B}} - p_j^{\mathcal{B}} \geq 0$

v $\theta_j^{\mathcal{B}} - p_j^{\mathcal{B}} \geq \theta_k^{\mathcal{B}} - p_k^{\mathcal{B}}, \forall k \neq j$

vi $\theta_j^{\mathcal{B}} - p_j^{\mathcal{B}} \geq \theta_k^{\mathcal{B}} + \eta^{\mathcal{B}} A_k, \forall k$

Define $W_j$ being the region on $\mathbb{R}^{J+1}$ defined by these constraints. Then, the proportion of consumers that pay for platform J is:

$$B_j = \int_{W_j} f_{\mathcal{B}}(\theta, \eta) d\theta \, d\eta \tag{3.20}$$

Proceeding as in Section 3.3.1, I can formulate the optimization problem of platform $j$ in term of quantities.

$$\begin{aligned} \max_{F^j, A^j} \Big\{ & - c_j^F \cdot F_j + (p^{B,j}(A_j, F_j, A_{-j}, F_{-j}) - c^{B,j}) \cdot B^j(A_j, F_j, A_{-j}, F_{-j}) \\ & + (p_j^A(A_j, F_j) - c_j^A) \cdot A_j \Big\} \end{aligned} \tag{3.21}$$

$$\text{s. t. } B(A_j, F_j, A_{-j}, F_{-j}) = \int_{W^j} f_{\mathcal{B}}(\theta, \eta) d\theta \, d\eta$$

Let quantities form a Nash equilibrium of the platform competition game. That is,

$$\forall j, F_j \text{ and } A_j \text{ maximize profits of platform } j \text{ given } F_{-j} \text{ and } A_{-j}$$

Define $\tilde{s}_j = -\frac{\partial F_j}{\partial B_j}$ and $\tilde{\eta}_j^A = \frac{\int_{-\infty}^{\infty} \eta \cdot f_{\mathcal{A}} \left( p_j^A - \eta_j F_j, \theta_{-j}, \eta \right) d\theta_{-j} \, d\eta}{\int_{-\infty}^{\infty} f_{\mathcal{A}} \left( p_j^A - \eta_j F_j, \theta_{-j}, \eta \right) d\theta_{-j} \, d\eta}$ the average interaction value of marginal users of side $\mathcal{A}$ from platform $j$. The following proposition characterizes the interior equilibrium under oligopolistic platform competition and extends Proposition 3.3.1 (J=1).

**Proposition 3.7.1.** *The profit-maximizing pricing of the oligopoly under freemium packages is given by:*

$$\frac{p_j^{\mathcal{A}} - c_j^{\mathcal{A}}}{p_j^{\mathcal{A}}} = \frac{1}{\epsilon_j^{\mathcal{A}}} - \frac{\epsilon_{F_j, B_j}}{J} \frac{p_j^{\mathcal{B}} F_j}{p_j^{\mathcal{A}} A_j}$$

$$\frac{p_j^{\mathcal{B}} - c_j^{\mathcal{B}}}{p_j^{\mathcal{B}}} = \frac{\epsilon_{F_j, B_j}}{\epsilon_j^{\mathcal{B}}} + \frac{\tilde{s}_j}{p_j^{\mathcal{B}}} \cdot [\tilde{\eta}_j^A A_j - c_j^F]$$

*Proof.* See appendix 3.8.3. □

## 3.8 Appendix IV: Proofs

### 3.8.1 Proof of Proposition 3.3.1

*Proof.* From the profit maximization of the platform:

$$-c^F + (p^{\mathcal{B}} - c^{\mathcal{B}})\frac{\partial B}{\partial F} + \frac{\partial p^{\mathcal{B}}}{\partial F}B + \frac{\partial p^{\mathcal{A}}}{\partial F}A = 0 \tag{3.22}$$

$$(p^{\mathcal{B}} - c^{\mathcal{B}})\frac{\partial B}{\partial A} + \frac{\partial p^{\mathcal{B}}}{\partial A}B + \frac{\partial p^{\mathcal{A}}}{\partial A}A + p^{\mathcal{A}} - c^{\mathcal{A}} = 0 \tag{3.23}$$

Using implicit function theorem on equations (3.1),(3.2),(3.3) we have:

From equation (3.1):

$$\frac{\partial p^{\mathcal{A}}}{\partial A} = \frac{-1}{\int_{-\infty}^{\infty} f_{\mathcal{A}}(p^{\mathcal{A}} - \eta F, \eta)d\eta} \tag{3.24}$$

$$\frac{\partial p^{\mathcal{A}}}{\partial F} = \frac{\int_{-\infty}^{\infty} \eta \cdot f_{\mathcal{A}}\Big(p^{\mathcal{A}} - \eta F, \eta\Big)d\eta}{\int_{-\infty}^{\infty} f_{\mathcal{A}}\Big(p^{\mathcal{A}} - \eta F, \eta\Big)d\eta} = \tilde{\eta}^{A} \tag{3.25}$$

where $\tilde{\eta}^A$ is the average interaction value of marginal users

From equation (3.2):

$$\frac{\partial p^{\mathcal{B}}}{\partial F} = \frac{A}{\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big)d\theta} \tag{3.26}$$

From equation (3.3):

$$\frac{\partial p^{\mathcal{B}}}{\partial A} = \frac{\frac{p^{\mathcal{B}}}{(A)^2} \cdot \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta}{\frac{1}{A}\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}(\theta, -\frac{p^{\mathcal{B}}}{A}) d\theta + \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} f_{\mathcal{B}}(p^{\mathcal{B}}, \eta) d\eta} \tag{3.27}$$

$$\begin{aligned}
\frac{\partial B}{\partial F} &= -\frac{\partial p^{\mathcal{B}}}{\partial F} \cdot \frac{1}{A} \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta - \frac{\partial p^{\mathcal{B}}}{\partial F} \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} f_{\mathcal{B}}\Big(p^{\mathcal{B}}, \eta\Big) d\eta \\
&= -1 - A \cdot \frac{\int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} f_{\mathcal{B}}\Big(p^{\mathcal{B}}, \eta\Big) d\eta}{\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta}
\end{aligned} \tag{3.28}$$

$$\begin{aligned}
\frac{\partial B}{\partial A} &= \Big[ -\frac{\partial p^{\mathcal{B}}}{\partial A} \cdot \frac{1}{A} + \frac{p^{\mathcal{B}}}{(A)^2} \Big] \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta - \frac{\partial p^{\mathcal{B}}}{\partial A} \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} f_{\mathcal{B}}\Big(p^{\mathcal{B}}, \eta\Big) d\eta \\
&= -\frac{\partial p^{\mathcal{B}}}{\partial A} \Big[ \frac{1}{A} \cdot \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta + \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} f_{\mathcal{B}}\Big(p^{\mathcal{B}}, \eta\Big) d\eta \Big] + \frac{p^{\mathcal{B}}}{(A)^2} \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta \\
&= -\frac{p^{\mathcal{B}}}{(A)^2} \int_{p^{\mathcal{A}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) dB + \frac{p^{\mathcal{B}}}{(A)^2} \int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta \\
&= 0
\end{aligned} \tag{3.29}$$

Denote $\tilde{s} = -\frac{\partial F}{\partial B}$ and $\mu^{\mathcal{I}} = \frac{p^{\mathcal{I}}}{\epsilon^{\mathcal{I}}}$. Using equations (3.24)-(3.29) in the F.O.C., I get the desired result:

$$\frac{p^{\mathcal{A}} - c^{\mathcal{A}}}{p^{\mathcal{A}}} = \frac{1}{\epsilon^{A}} - \epsilon_{F,B} \cdot \frac{p^{\mathcal{B}} \cdot F}{p^{\mathcal{A}} \cdot A}$$

$$\frac{p^{\mathcal{B}} - c^{\mathcal{B}}}{p^{\mathcal{B}}} = \frac{\epsilon_{F,B}}{\epsilon^{B}} + \frac{\tilde{s}}{p^{\mathcal{B}}} \cdot [\tilde{\eta}^{A} A - c^{F}]$$

□

### 3.8.2 Proof of Proposition 3.3.2

*Proof.* From the Social Planner's problem:

$$\frac{\partial V^A}{\partial F} + \frac{\partial V^F}{\partial F} + \frac{\partial V^B}{\partial F} - c^F - c^{\mathcal{B}} \frac{\partial B}{\partial F} = 0 \tag{3.30}$$

$$\frac{\partial V^A}{\partial A} + \frac{\partial V^F}{\partial A} + \frac{\partial V^B}{\partial A} - c^{\mathcal{A}} - c^{\mathcal{B}} \frac{\partial B}{\partial A} = 0 \tag{3.31}$$

Using Leibniz rule and implicit function theorem:

From equation (3.8):

$$\begin{aligned}
\frac{\partial V^A}{\partial F} &= \overbrace{-p^{\mathcal{A}} \cdot \int_{\infty}^{\infty} \Big(\frac{\partial p^{\mathcal{A}}}{\partial F} - \eta\Big) f_{\mathcal{A}}(p^{\mathcal{A}} - \eta F, \eta) d\eta}^{=0,\text{ by (3.25)}} + \int_{-\infty}^{\infty} \int_{p^{\mathcal{A}} - \eta F}^{\infty} \eta f_{\mathcal{A}}(\theta, \eta) d\theta \, d\eta \\
&= \int_{-\infty}^{\infty} \int_{p^{\mathcal{A}} - \eta F}^{\infty} \eta f_{\mathcal{A}}(\theta, \eta) d\theta \, d\eta \\
&= \bar{\eta}^A A
\end{aligned} \tag{3.32}$$

where $\bar{\eta}^A$ is the average interaction value of average users from side $\mathcal{A}$.

From equations (3.9) and (3.26):

$$\begin{aligned}
\frac{\partial V^F}{\partial F} &= \frac{\partial p^{\mathcal{B}}}{\partial F} \frac{1}{A} \int_{p^{\mathcal{B}}}^{\infty} (\theta - p^{\mathcal{B}}) f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta \\
&= \frac{1}{\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta} \cdot \int_{p^{\mathcal{B}}}^{\infty} (\theta - p^{\mathcal{B}}) f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta \\
&= \frac{\int_{p^{\mathcal{B}}}^{\infty} \theta f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta}{\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta, -\frac{p^{\mathcal{B}}}{A}\Big) d\theta} - p^{\mathcal{B}} \\
&= \bar{\theta}^B - p^{\mathcal{B}}
\end{aligned} \tag{3.33}$$

where $\bar{\theta}^B$ is the average membership value of users who pay that are on the margin between using the platform for free and their outside option.

From equations (3.10), (3.26) and (3.28):

$$\begin{aligned}
\frac{\partial V^B}{\partial F} &= -\frac{\partial p^{\mathcal{B}}}{\partial F}\Big[\frac{1}{A}\int_{p^{\mathcal{B}}}^{\infty}\theta f_{\mathcal{B}}\Big(\theta,-\frac{p^{\mathcal{B}}}{A}\Big)d\theta + \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} p^{\mathcal{B}} f_{\mathcal{B}}(p^{\mathcal{B}},\eta)d\eta\Big] \\
&= -\frac{A}{\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta,-\frac{p^{\mathcal{B}}}{A}\Big)d\theta}\Big[\frac{1}{A}\int_{p^{\mathcal{B}}}^{\infty}\theta f_{\mathcal{B}}\Big(\theta,-\frac{p^{\mathcal{B}}}{A}\Big)d\theta + p^{\mathcal{B}}\int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} f_{\mathcal{B}}(p^{\mathcal{B}},\eta)d\eta\Big] \\
&= -\bar{\theta}^B + \Big(\frac{\partial B}{\partial F}+1\Big)p^{\mathcal{B}} \qquad (3.34)
\end{aligned}$$

Now,

From equations (3.8) and (3.24):

$$\frac{\partial V^A}{\partial A} = -p^{\mathcal{A}}\cdot\frac{\partial p^{\mathcal{A}}}{\partial A}\cdot\int_{-\infty}^{\infty} f_{\mathcal{A}}(p^{\mathcal{A}}-\eta F,\eta)d\eta = p^{\mathcal{A}} \qquad (3.35)$$

From equation (3.9):

$$\frac{\partial V^F}{\partial A} = \Big[\frac{\partial p^{\mathcal{B}}}{\partial A}\cdot\frac{1}{A} - \frac{p^{\mathcal{B}}}{(A)^2}\Big]\cdot\int_{p^{\mathcal{B}}}^{\infty}(\theta-p^{\mathcal{B}})f_{\mathcal{B}}\Big(\theta,-\frac{p^{\mathcal{B}}}{A}\Big)\,d\theta \qquad (3.36)$$

From equation (3.10):

$$\begin{aligned}
\frac{\partial V^B}{\partial A} = \Big[-\frac{\partial p^{\mathcal{B}}}{\partial A}\cdot\frac{1}{A} + \frac{p^{\mathcal{B}}}{(A)^2}\Big]\cdot\int_{p^{\mathcal{B}}}^{\infty}\theta f_{\mathcal{B}}\Big(\theta,-\frac{p^{\mathcal{B}}}{A}\Big)\,d\theta \\
-\frac{\partial p^{\mathcal{B}}}{\partial A}\int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} p^{\mathcal{B}} f_{\mathcal{B}}(p^{\mathcal{B}},\eta)d\,\eta \qquad (3.37)
\end{aligned}$$

Note that:

$$\begin{aligned}
&\frac{\partial V^F}{\partial A} + \frac{\partial V^B}{\partial A} = \\
&= \Big[-\tfrac{\partial p^{\mathcal{B}}}{\partial A}\cdot\tfrac{1}{A} + \tfrac{p^{\mathcal{B}}}{(A)^2}\Big]\cdot p^{\mathcal{B}}\textstyle\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta,-\tfrac{p^{\mathcal{B}}}{A}\Big)\,d\theta - \tfrac{\partial p^{\mathcal{B}}}{\partial A}\int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} p^{\mathcal{B}} f_{\mathcal{B}}(p^{\mathcal{B}},\eta)d\,\eta \\
&= -p^{\mathcal{B}}\cdot\tfrac{\partial p^{\mathcal{B}}}{\partial A}\Big[\tfrac{1}{A}\textstyle\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta,-\tfrac{p^{\mathcal{B}}}{A}\Big)\,d\theta + \int_{-\infty}^{-\frac{p^{\mathcal{B}}}{A}} f_{\mathcal{B}}(p^{\mathcal{B}},\eta)d\,\eta\Big] + \Big(\tfrac{p^{\mathcal{B}}}{A}\Big)^2\textstyle\int_{p^{\mathcal{B}}}^{\infty} f_{\mathcal{B}}\Big(\theta,-\tfrac{p^{\mathcal{B}}}{A}\Big)\,d\theta \\
&=^{(3.27)} 0 \qquad (3.38)
\end{aligned}$$

Substituting equations (3.32) (3.38) into the above mentioned F.O.C.'s we get:

$$p^{\mathcal{B}} = c^{\mathcal{B}} - \tilde{s} \cdot c^F + \tilde{s} \cdot \bar{\eta}^A A$$

$$p^{\mathcal{A}} = c^{\mathcal{A}}$$

□

### 3.8.3 Proof of Proposition 3.7.1

*Proof.* Note that

$$\begin{aligned} V_j = \Big\{(\theta_1, ..., \theta_J, \eta) \in \mathbb{R}^{J+1} | \theta_j \geq -\eta_j A_j;& \\ \eta_j \geq -\frac{p_j^B}{A_j}& \\ \theta_k \leq \theta_j + \eta_j A_j + p_k^{\mathcal{B}}, \forall k \neq j;& \\ \eta_k \leq \frac{1}{A_k} \cdot \Big(\theta_j - \theta_k + \eta_j A_j\Big), \forall k \neq j\Big\}& \end{aligned}$$

$$\begin{aligned} W_j = \{(\theta_1, ..., \theta_J, \eta) \in \mathbb{R}^{J+1} | \theta_j \geq p_j^B;& \\ \eta_j \leq -\frac{p_j^B}{A_j}& \\ \theta_k \leq \theta_j - p_j^{\mathcal{B}} + p_k^{\mathcal{B}}, \forall k \neq j;& \\ \eta_k \leq \frac{1}{A_k} \cdot \Big(\theta_j - \theta_k - p_j^{\mathcal{B}}\Big), \forall k \neq j\Big\}& \end{aligned}$$

Define $\theta_{-j} = (\theta_1, ...\theta_{j-1}, \theta_{j+1}, ..., \theta_J)$ and $\eta_{-j} = (\eta_1, ...\eta_{j-1}, \eta_{j+1}, ..., \eta_J)$, and as an abuse of notation,

$$\theta_j + \eta_j A_j + p_{-j}^{\mathcal{B}} = \theta_j + \eta_j A_j + (p_1^{\mathcal{B}}, ..., p_{j-1}^{\mathcal{B}}, p_{j+1}^{\mathcal{B}}, ..., p_J^{\mathcal{B}})$$

$$\theta_j - \theta_{-j} + \eta_j A_j = \theta_j + \eta_j A_j - (\theta_1, ..., \theta_{j-1}, \theta_{j+1}, ..., \theta_J)$$

Then,

$$F_j = \int_{-\frac{p_j^{\mathcal{B}}}{A_j}}^{\infty} d\eta_j \int_{-\eta_j A_j}^{\infty} d\theta_j \int_{-\infty}^{\theta_j+\eta_j A_j+p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j-\theta_{-j}+\eta_j A_j)} d\eta_{-j} f_{\mathcal{B}}(\theta,\eta) \tag{3.39}$$

Similarly,

$$B_j = \int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j-p_j^{\mathcal{B}}+p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j-\theta_{-j}-p_j^{\mathcal{B}})} d\eta_{-j} f_{\mathcal{B}}(\theta,\eta) \tag{3.40}$$

Moreover, as in (3.18):

$$A_j = \int_{-\infty}^{\infty}\int_{p_j^{\mathcal{A}}-\eta_j F_j}^{\infty} f_{\mathcal{A}}(\theta,\eta) d\theta_j\, d\theta_{-j} d\eta \tag{3.41}$$

Now, note that $\forall j \in \{1,...,J\}$:

$$\frac{\partial A_j}{\partial p_j^{\mathcal{A}}} = -\int_{-\infty}^{\infty} f_{\mathcal{A}}(\overbrace{p_j^{\mathcal{A}} - \eta_j F_j}^{\theta_j}, \theta_{-j}, \eta) d\theta_{-j}\, d\eta < 0 \tag{3.42}$$

$$\frac{\partial F_j}{\partial p_j^{\mathcal{B}}} = \frac{1}{A_j}\int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j-p_j^{\mathcal{B}}+p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j-\theta_{-j}-p_j^{\mathcal{B}})} d\eta_{-j} f_{\mathcal{B}}\Big(\theta, \overbrace{-\frac{p_j^{\mathcal{B}}}{A_j}}^{\eta_j}, \eta_{-j}\Big) > 0 \tag{3.43}$$

Thus, fixing $F_j$, we can invert (3.42) as $p_j^{\mathcal{A}} = p^{\mathcal{A}}(A_j, F_j)$. Also, fixing $A_k > 0, \forall k$ and $F_{-j}$, we can invert equation (3.43) as $p^{\mathcal{B}} = p^{\mathcal{B}}(A_j, A_{-j}, F_j, F_{-j})$. Moreover, by equation (3.40), $B_j$ is uniquely pinned down.

Now we are ready to analyze the profit maximization problem of platform $j$:

$$-c_j^F + (p_j^B - c_j^B)\frac{\partial B_j}{\partial F_j} + \frac{\partial p_j^B}{\partial F_j}B_j + \frac{\partial p_j^A}{\partial F_j}A_j = 0 \tag{3.44}$$

$$(p_j^B - c_j^B)\frac{\partial B_j}{\partial A_j} + \frac{\partial p_j^B}{\partial A_j}B_j + \frac{\partial p_j^A}{\partial A_j}A_j + p_j^A - c_j^A = 0 \tag{3.45}$$

From equation (3.41)

$$\frac{\partial p_j^{\mathcal{A}}}{\partial A_j} = -\frac{1}{\int_{-\infty}^{\infty} f_{\mathcal{A}}(\underbrace{p_j^{\mathcal{A}} - \eta_j F_j}_{\theta_j}, \theta_{-j}, \eta) d\theta_{-j}\, d\eta} \tag{3.46}$$

$$\frac{\partial p_j^{\mathcal{A}}}{\partial F_j} = \frac{\int_{-\infty}^{\infty} \eta \cdot f_{\mathcal{A}}\Big(p_j^{\mathcal{A}} - \eta_j F_j, \theta_{-j}, \eta\Big) d\theta_{-j}\, d\eta}{\int_{-\infty}^{\infty} f_{\mathcal{A}}\Big(p_j^{\mathcal{A}} - \eta_j F_j, \theta_{-j}, \eta\Big) d\theta_{-j}\, d\eta} = \tilde{\eta}_j^A \tag{3.47}$$

where $\tilde{\eta}_j^A$ is the average interaction value of users on side $\mathcal{A}$ that are on the margin from entering or not on the platform $j$.

From equation (3.39)

$$\frac{\partial p_j^{\mathcal{B}}}{\partial F_j} = \frac{A_j}{\int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j} f_{\mathcal{B}}\Big(\theta, \underbrace{-\frac{p_j^{\mathcal{B}}}{A_j}}_{\eta_j}, \eta_{-j}\Big)} \tag{3.48}$$

From equation (3.40)

$$\frac{\partial p_j^{\mathcal{B}}}{\partial A_j} = \frac{\frac{p_j^{\mathcal{B}}}{(A_j)^2} \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j} f_{\mathcal{B}}\Big(\theta, -\frac{p_j^{\mathcal{B}}}{A_j}, \eta_{-j}\Big)}{\Omega_j} \tag{3.49}$$

where

$$\Omega_j = \frac{1}{A_j}\int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j} f_{\mathcal{B}}\Big(\theta, -\frac{p_j^{\mathcal{B}}}{A_j}, \eta_{-j}\Big)$$
$$+ \int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{-\infty}^{p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{-\frac{\theta_{-j}}{A_{-j}}} d\eta_{-j} f_{\mathcal{B}}(p_j^{\mathcal{B}}, \theta_{-j}, \eta)$$
$$+ \textstyle\sum_{k\neq j} \int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{-\infty}^{-\frac{p_k^{\mathcal{B}}}{A_k}} d\eta_k \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j,-k}^{\mathcal{B}}} d\theta_{-j,-k} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j,-k} f_{\mathcal{B}}(\underbrace{\theta_j - p_j^{\mathcal{B}} + p_k^{\mathcal{B}}}_{\theta_k}, \theta_{-k}, \eta)$$
$$+ \textstyle\sum_{k\neq j} \frac{1}{A_k} \int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j,-k}}(\theta_j - \theta_{-j,-k} - p_j^{\mathcal{B}})} d\eta_{-j,-k} f_{\mathcal{B}}\Big(\theta, \underbrace{\frac{1}{A_k}(\theta_j - \theta_k - p_j^{\mathcal{B}})}_{\eta_k}, \eta_{-k}\Big) \tag{3.50}$$

$$\frac{\partial B_j}{\partial A_j} =$$
$$= -\frac{\partial p_j^{\mathcal{B}}}{\partial A_j}\Omega_j + \frac{p_j^{\mathcal{B}}}{(A_j)^2}\int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j}^{\mathcal{B}}} d\theta_{-j} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j} f_{\mathcal{B}}\Big(\theta, -\frac{p_j^{\mathcal{B}}}{A_j}, \eta_{-j}\Big)$$
$$+ \textstyle\sum_{k\neq j} \frac{\partial p_k^{\mathcal{B}}}{\partial A_j} \int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{-\infty}^{-\frac{p_k^{\mathcal{B}}}{A_k}} d\eta_k \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j,-k}^{\mathcal{B}}} d\theta_{-j,-k} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j,-k} f_{\mathcal{B}}(\theta_j - p_j^{\mathcal{B}} + p_k^{\mathcal{B}}, \theta_{-k}, \eta)$$
$$=^{by (3.49)} \textstyle\sum_{k\neq j} \frac{\partial p_k^{\mathcal{B}}}{\partial A_j} \int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{-\infty}^{-\frac{p_k^{\mathcal{B}}}{A_k}} d\eta_k \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j,-k}^{\mathcal{B}}} d\theta_{-j,-k} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j,-k} f_{\mathcal{B}}(\theta_j - p_j^{\mathcal{B}} + p_k^{\mathcal{B}}, \theta_{-k}, \eta) \tag{3.51}$$

$$\frac{\partial p_k^{\mathcal{B}}}{\partial F_j} = -\frac{\partial B_j/\partial F_j}{\partial B_j/\partial p_k^{\mathcal{B}}}$$
$$= \frac{-\partial B_j/\partial F_j}{\int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{-\infty}^{-\frac{p_k^{\mathcal{B}}}{A_k}} d\eta_k \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j,-k}^{\mathcal{B}}} d\theta_{-j,-k} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j,-k} f_{\mathcal{B}}(\theta_j - p_j^{\mathcal{B}} + p_k^{\mathcal{B}}, \theta_{-k}, \eta)} \tag{3.52}$$

$$\frac{\partial p_k^{\mathcal{B}}}{\partial A_j} = -\frac{\partial B_j/\partial A_j}{\partial B_j/\partial p_k^{\mathcal{B}}}$$
$$= \frac{-\partial B_j/\partial A_j}{\int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{-\infty}^{-\frac{p_k^{\mathcal{B}}}{A_k}} d\eta_k \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j,-k}^{\mathcal{B}}} d\theta_{-j,-k} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j,-k} f_{\mathcal{B}}(\theta_j - p_j^{\mathcal{B}} + p_k^{\mathcal{B}}, \theta_{-k}, \eta)} \tag{3.53}$$

$$\begin{aligned}
\frac{\partial B_j}{\partial F_j} &= -\frac{\partial p_j^{\mathcal{B}}}{\partial F_j}\Omega_j \\
&+ \textstyle\sum_{k\neq j} \frac{\partial p_k^{\mathcal{B}}}{\partial F_j} \int_{-\infty}^{-\frac{p_j^{\mathcal{B}}}{A_j}} d\eta_j \int_{-\infty}^{-\frac{p_k^{\mathcal{B}}}{A_k}} d\eta_k \int_{p_j^{\mathcal{B}}}^{\infty} d\theta_j \int_{-\infty}^{\theta_j - p_j^{\mathcal{B}} + p_{-j,-k}^{\mathcal{B}}} d\theta_{-j,-k} \int_{-\infty}^{\frac{1}{A_{-j}}(\theta_j - \theta_{-j} - p_j^{\mathcal{B}})} d\eta_{-j,-k} f_{\mathcal{B}}(\theta_j - p_j^{\mathcal{B}} + p_k^{\mathcal{B}}, \theta_{-k}, \eta) \\
&=^{(3.52)} -\frac{\partial p_j^{\mathcal{B}}}{\partial F_j}\Omega_j - (J-1)\frac{\partial B_j}{\partial F_j}
\end{aligned}$$

Thus, an important relation is:

$$\frac{\partial B_j}{\partial F_j} = -\frac{\partial p_j^{\mathcal{B}}}{\partial F_j} \cdot \frac{\Omega_j}{J} \tag{3.54}$$

Defining $\tilde{s}_j = -\frac{\partial F_j}{\partial B_j}$ it follows from (3.48), (3.49) and (3.54) that

$$\frac{\partial p_j^{\mathcal{B}}}{\partial A_j} = \frac{\tilde{s}_j}{J} \cdot \frac{p_j^{\mathcal{B}}}{A_j} \tag{3.55}$$

Moreover, from (3.51) and (3.53):

$$\frac{\partial B_j}{\partial A_j} = 0 \tag{3.56}$$

Substituting these equations into the F.O.C.'s I get:

$$\frac{p_j^A - c_j^{\mathcal{A}}}{p_j^{\mathcal{A}}} = \frac{1}{\epsilon_j^A} - \frac{\epsilon_{F_j,B_j}}{J} \frac{p_j^{\mathcal{B}} F_j}{p_j^{\mathcal{A}} A_j}$$

$$\frac{p_j^{\mathcal{B}} - c_j^{\mathcal{B}}}{p_j^{\mathcal{B}}} = \frac{\epsilon_{F_j,B_j}}{\epsilon_j^B} + \frac{\tilde{s}_j}{p_j^{\mathcal{B}}} \cdot [\tilde{\eta}_j^A A_j - c_j^F]$$

$\square$

### 3.8.4 Proof of Proposition 3.5.1

*Proof.*

$$\begin{aligned}
B(x_i, p^{\mathcal{B}}, A) &= Pr(X \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} \geq 0; \quad X \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} \geq X \cdot \beta - \eta \cdot A + u^{\mathcal{B}} | X = x_i) \\
&= Pr(x_i \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} \geq 0; \quad \eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \qquad (3.57) \\
&=^{x_i^1 \to \infty} Pr(\eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \\
&= 1 - F_{\eta | X = x_i}\Big(\frac{p^{\mathcal{B}}}{A}\Big) \\
&= 1 - F_{\eta}\Big(\frac{p^{\mathcal{B}}}{A}\Big) \qquad (3.58)
\end{aligned}$$

Thus,

$$F_{\eta}\Big(\frac{p^{\mathcal{B}}}{A}\Big) = 1 - B(x_i, p^{\mathcal{B}}, A) \qquad (3.59)$$

And $F_{\eta}(\cdot)$ is identified on the support of $\frac{p^{\mathcal{B}}}{A}$, when $x_i^1 \to \infty$. Since the distribution of $X$ is known, $F_{\eta}$ is identified.

Now, going back to (3.57):

$$\begin{aligned}
B(x_i, p^{\mathcal{B}}, A) &= Pr(x_i \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} \geq 0; \quad \eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \\
&= Pr(u^{\mathcal{B}} \geq p^{\mathcal{B}} - x_i \cdot \beta; \quad \eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \\
&= Pr(u^{\mathcal{B}} \geq p^{\mathcal{B}} - x_i \cdot \beta | X = x_i) P(\eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \\
&= \Big[1 - F_{u|X=x_i}(p^{\mathcal{B}} - x_i \cdot \beta)\Big]\Big[1 - F_{\eta|X=x_i}\Big(\frac{p^{\mathcal{B}}}{A}\Big)\Big] \\
&= \Big[1 - F_u(p^{\mathcal{B}} - x_i \cdot \beta)\Big]\Big[1 - F_{\eta}\Big(\frac{p^{\mathcal{B}}}{A}\Big)\Big] \qquad (3.60)
\end{aligned}$$

Thus,

$$F_u(p^{\mathcal{B}} - x_i \cdot \beta) = 1 - \frac{B(x_i, p^{\mathcal{B}}, A)}{1 - F_{\eta}\Big(\frac{p^{\mathcal{B}}}{A}\Big)} \qquad (3.61)$$

Now, note that equation (3.61) resembles a single index model. On the right everything is known, while on the left of (3.61) $F_u$ and $\beta$ are unknown. Utilizing the normalization assumption that $||\beta||= 1$ and the support assumptions on $X$, I can first identify $\beta$, and then $F_u$, as follows:

$$F_u(p^{\mathcal{B}} - x_i \cdot \beta) = F_{\tilde{u}}(p^{\mathcal{B}} - x_i \cdot \tilde{\beta}) \Rightarrow x_i \cdot \beta = p^{\mathcal{B}} - F_u^{-1}\Big(F_{\tilde{u}}(p^{\mathcal{B}} - x_i \cdot \tilde{\beta})\Big)$$
$$\Rightarrow x_i \cdot \beta = a + b(x_i \cdot \tilde{\beta}) \tag{3.62}$$

Plugging values from the support of $X$:

$$x_i = 0 \Rightarrow a = 0 \tag{3.63}$$

$$x_i = e_k \Rightarrow \beta_k = b \cdot \tilde{\beta}_k \Rightarrow \beta = b \cdot \tilde{\beta} \Rightarrow ||\beta||= |b|\cdot||\tilde{\beta}||\Rightarrow |b|= 1 \tag{3.64}$$

Since $p^{\mathcal{B}} - F_u^{-1}\Big(F_{\tilde{u}}(p^{\mathcal{B}} - x_i \cdot \tilde{\beta})\Big)$ is strictly increasing in $x_i$, it has to be that $b = 1$. Thus, from (3.64) we have that $\beta = \tilde{\beta}$, which shows that $\beta$ is identified.

Going back to (3.56), $F_u$ is identified on the support of $p^{\mathcal{B}} - x_i \cdot \beta$. □

### 3.8.5 Proof of Proposition 3.5.2

*Proof.* On one hand,

$$\begin{aligned}
B(x_i, p^{\mathcal{B}}, A) &= Pr(X \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} \geq 0; \quad X \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} \geq X \cdot \beta - \eta \cdot A + u^{\mathcal{B}} | X = x_i) \\
&= Pr(x_i \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} \geq 0; \quad \eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \\
&= Pr(u^{\mathcal{B}} \geq p^{\mathcal{B}} - x_i \cdot \beta; \quad \eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \\
&= Pr(u^{\mathcal{B}} \geq p^{\mathcal{B}} - x_i \cdot \beta | X = x_i) P(\eta \cdot A \geq p^{\mathcal{B}} | X = x_i) \\
&= \Big[1 - F_{u^{\mathcal{B}}|X=x_i}(p^{\mathcal{B}} - x_i \cdot \beta)\Big]\Big[1 - F_{\eta|X=x_i}\Big(\frac{p^{\mathcal{B}}}{A}\Big)\Big] \\
&= \Big[1 - F_{u^{\mathcal{B}}}(p^{\mathcal{B}} - x_i \cdot \beta)\Big]\Big[1 - F_{\eta}\Big(\frac{p^{\mathcal{B}}}{A}\Big)\Big]
\end{aligned} \tag{3.65}$$

On the other hand,

$$
\begin{aligned}
F(x_i, p^{\mathcal{B}}, A) &= Pr(X_i \cdot \beta - \eta \cdot A + u^{\mathcal{B}} \geq 0; \quad X_i \cdot \beta - \eta \cdot A + u^{\mathcal{B}} \geq X_i \cdot \beta - p^{\mathcal{B}} + u^{\mathcal{B}} | X = x_i) \\
&= Pr(x_i \cdot \beta - \eta \cdot A + u^{\mathcal{B}} \geq 0; \quad \eta \cdot A \leq p^{\mathcal{B}} | X = x_i) \\
&= Pr(u^{\mathcal{B}} \geq \eta \cdot A - x_i \cdot \beta; \quad \eta \cdot A \leq p^{\mathcal{B}} | X = x_i) \\
&= \int_{-\infty}^{p^{\mathcal{B}}/A} \int_{\eta \cdot A - x_i \cdot \beta}^{\infty} f_{u^{\mathcal{B}}|X=x_i}(u) \cdot f_{\eta|X=x_i}(\eta) du \, d\eta \\
&= \int_{-\infty}^{p^{\mathcal{B}}/A} \Big[1 - F_{u^{\mathcal{B}}|X=x_i}(\eta \cdot A - x_i \cdot \beta)\Big] \cdot f_{\eta|X=x_i}(\eta) d\eta \\
&= \int_{-\infty}^{p^{\mathcal{B}}/A} \Big[1 - F_{u^{\mathcal{B}}}(\eta \cdot A - x_i \cdot \beta)\Big] \cdot f_{\eta}(\eta) d\eta
\end{aligned}
\tag{3.66}
$$

From (3.65),

$$
\begin{aligned}
B(x_i, p^{\mathcal{B}}, A) &= \Big[1 - F_{u^{\mathcal{B}}}(p^{\mathcal{B}} - x_i \cdot \beta)\Big]\Big[1 - F_{\eta}\Big(\frac{p^{\mathcal{B}}}{A}\Big)\Big] \\
\Rightarrow \Big[1 - F_{u^{\mathcal{B}}}(\eta A - x_i \cdot \beta)\Big] &= \frac{B(x_i, \eta A, A)}{1 - F_{\eta}(\eta)}
\end{aligned}
\tag{3.67}
$$

Substituting (3.67) into equation (3.66):

$$
F(x_i, p^{\mathcal{B}}, A) = \int_{-\infty}^{p^{\mathcal{B}}/A} B(x_i, \eta A, A) \frac{f_{\eta}(\eta)}{1 - F_{\eta}(\eta)} d\eta \tag{3.68}
$$

Thus,

$$
\begin{aligned}
\frac{\partial F(x_i, p^{\mathcal{B}}, A)}{\partial p^{\mathcal{B}}} &= \frac{1}{A} B(x_i, p^{\mathcal{B}}, A) \frac{f_\eta\left(\frac{p^{\mathcal{B}}}{A}\right)}{1 - F_\eta\left(\frac{p^{\mathcal{B}}}{A}\right)} \\
\Rightarrow \frac{f_\eta\left(\frac{p^{\mathcal{B}}}{A}\right)}{1 - F_\eta\left(\frac{p^{\mathcal{B}}}{A}\right)} &= \frac{A \cdot \frac{\partial F(x_i, p^{\mathcal{B}}, A)}{\partial p^{\mathcal{B}}}}{B(x_i, p^{\mathcal{B}}, A)} \\
\Rightarrow -\frac{\partial log\left(1 - F_\eta\left(\frac{p^{\mathcal{B}}}{A}\right)\right)}{\partial p^{\mathcal{B}}} &= \frac{\frac{\partial F(x_i, p^{\mathcal{B}}, A)}{\partial p^{\mathcal{B}}}}{B(x_i, p^{\mathcal{B}}, A)} \\
\Rightarrow log\Big(1 - F_\eta(\eta)\Big) &= -\int_{-\infty}^{\eta A} \frac{\frac{\partial F(x_i, p^{\mathcal{B}}, A)}{\partial p^{\mathcal{B}}}}{B(x_i, p^{\mathcal{B}}, A)} dp^{\mathcal{B}} \\
\Rightarrow F_\eta(\eta) = 1 - \exp &\left\{ -\int_{-\infty}^{\eta A} \frac{\frac{\partial F(x_i, p^{\mathcal{B}}, A)}{\partial p^{\mathcal{B}}}}{B(x_i, p^{\mathcal{B}}, A)} dp^{\mathcal{B}} \right\} \qquad (3.69)
\end{aligned}
$$

Thus, $F_\eta$ is identified.

Finally, going back to (3.65):

$$
F_{u^{\mathcal{B}}}(p^{\mathcal{B}} - x_i \cdot \beta) = 1 - B(x_i, p^{\mathcal{B}}, A) \cdot \exp\left\{ \int_{-\infty}^{p^{\mathcal{B}}} \frac{\frac{\partial F(x_i, s, A)}{\partial s}}{B(x_i, s, A)} ds \right\} \qquad (3.70)
$$

By the same argument used in Proposition 3.5.1, $\beta$ is identified. With enough variation on prices $p^{\mathcal{B}}$, $F_{u^{\mathcal{B}}}$ is identified.

□

### 3.8.6 Proof of Proposition 3.5.3

*Proof.*

$$\begin{aligned}
A(y_i) &= Pr(-Y \cdot \xi + \eta \cdot F - p^{\mathcal{A}} + u^{\mathcal{A}} \geq 0 | Y = y_i) \\
&= Pr(u^{\mathcal{A}} + \eta \cdot F \geq y_i \cdot \xi + p^{\mathcal{A}} | Y = y_i) \\
&= \int_{y_i \cdot \xi + p^{\mathcal{A}}} f_{u^{\mathcal{A}} + \eta \cdot F | Y = y_i}(e) de \\
&= 1 - F_{u^{\mathcal{A}} + \eta \cdot F | Y = y_i}(y_i \cdot \xi + p^{\mathcal{A}}) \\
&= 1 - F_{u^{\mathcal{A}} + \eta \cdot F}(y_i \cdot \xi + p^{\mathcal{A}}) \qquad (3.71)
\end{aligned}$$

Thus,

$$F_{u^{\mathcal{A}} + \eta \cdot F}(y_i \cdot \xi + p^{\mathcal{A}}) = 1 - A(y_i) \qquad (3.72)$$

By the same argument used in Proposition 3.5.1, $\xi$ is identified.

Moreover, with enough variation on prices $p^{\mathcal{A}}$, $F_{u^{\mathcal{A}} + \eta \cdot F}$ is identified.

Now, denoting by $M_X(t)$ the moment generating function of the random variable $X$ at the point $t$ and utilizing that $u^{\mathcal{A}} \perp \eta$:

$$\begin{aligned}
M_{u^{\mathcal{A}} + \eta \cdot F}(t) &= M_{u^{\mathcal{A}}}(t) \cdot M_{\eta \cdot F}(t) \\
&= M_{u^{\mathcal{A}}}(t) \cdot M_{\eta}(t \cdot F) \qquad (3.73)
\end{aligned}$$

So that,

$$M_{u^{\mathcal{A}}}(t) = \frac{\overbrace{M_{u^{\mathcal{A}} + \eta \cdot F}(t)}^{\text{known by (20)}}}{\underbrace{M_{\eta}(t \cdot F)}_{\text{known by Prop. 3.5.2}}}$$

Since two distributions that have the same moment-generating function are identical at almost all points, $F_{u^{\mathcal{A}}}$ is identified. $\square$

# Bibliography

[1] Angelucci, C., Cage, J., and de Nijs, R. (2013). Price Discrimination in a Two-Sided Market: Theory and Evidence from the Newspaper Industry. Working Papers 13-13, NET Institute.

[2] Argentesi, E. and Filistrucchi, L. (2007). Estimating market power in a two-sided market: The case of newspapers. *Journal of Applied Econometrics*, 22(7):1247–1266.

[3] Armstrong, M. (2006). Competition in two-sided markets. *RAND Journal of Economics*, 37(3):668–691.

[4] Asker, J. (2010). A study of the internal organization of a bidding cartel. *American Economic Review*, 100(3):724–62.

[5] Athey, S., Coey, D., and Levin, J. (2013). Set-asides and subsidies in auctions. *American Economic Journal: Microeconomics*, 5(1):1–27.

[6] Athey, S. and Haile, P. A. (2002). Identification of standard auction models. *Econometrica*, 70(6):2107–2140.

[7] Athey, S., Levin, J., and Seira, E. (2011). Comparing open and Sealed Bid Auctions: Evidence from Timber Auctions*. *The Quarterly Journal of Economics*, 126(1):207–257.

[8] Bajari, P., Hong, H., and Ryan, S. P. (2010). Identification and estimation of a discrete game of complete information. *Econometrica*, 78(5):1529–1568.

[9] Bajari, P. and Hortaçsu, A. (2003). The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *The RAND Journal of Economics*, 34(2):329–355.

[10] Beresteanu, A., Molchanov, I., and Molinari, F. (2011). Sharp Identification Regions in Models With Convex Moment Predictions. *Econometrica*, 79(6):1785–1821.

[11] Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, 63(4):841–90.

[12] Berry, S. and Tamer, E. (2007). Identification in Models of Oligopoly Entry. *Advances in Economics and Econometrics: Theory and Applications*, II(4):46–85.

[13] Berry, S. T. (1990). Airport Presence as Product Differentiation. *American Economic Review*, 80(2):394–99.

[14] Berry, S. T. (1992a). Estimation of a Model of Entry in the Airline Industry. *Econometrica*, 60(4):889–917.

[15] Berry, S. T. (1992b). Estimation of a model of entry in the airline industry. *Econometrica*, 60(4):889–917.

[16] Bjorn, P. A. and Vuong, Q. H. (1984). Simultaneous Equations Models for Dummy Endogenous Variables: A Game Theoretic Formulation with an Application to Labor Force Participation. Working Papers 537, California Institute of Technology, Division of the Humanities and Social Sciences.

[17] Brendstrup, B. and Paarsch, H. J. (2006). Identification and estimation in sequential, asymmetric, english auctions. *Journal of Econometrics*, 134(1):69 – 94.

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

[19] Bulow, J. and Klemperer, P. (2009). Why do sellers (usually) prefer auctions? *American Economic Review*, 99(4):1544–75.

[20] Busse, M. and Rysman, M. (2005). Competition and Price Discrimination in Yellow Pages Advertising. *RAND Journal of Economics*, 36(2):378–390.

[21] Böhme, E. (2014). Second-Degree Price Discrimination on Two-Sided Markets. Annual Conference 2014 (Hamburg): Evidence-based Economic Policy 100507, Verein für Socialpolitik / German Economic Association.

[22] Caillaud, B. and Jullien, B. (2003). Chicken & Egg: Competition among Intermediation Service Providers. *RAND Journal of Economics*, 34(2):309–28.

[23] Chandra, A. and Collard-Wexler, A. (2009). Mergers in Two-Sided Markets: An Application to the Canadian Newspaper Industry. *Journal of Economics & Management Strategy*, 18(4):1045–1070.

[24] Chernozhukov, V., Lee, S., and Rosen, A. M. (2013). Intersection bounds: Estimation and inference. *Econometrica*, 81(2):667–737.

[25] Ciliberto, F. and Tamer, E. (2009a). Market Structure and Multiple Equilibria in Airline Markets. *Econometrica*, 77(6):1791–1828.

[26] Ciliberto, F. and Tamer, E. (2009b). Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828.

[27] Ferraz, C., Finan, F., and Szerman, D. (2015). Procuring firm growth: The effects of government purchases on firm dynamics. Working Paper 21219, National Bureau of Economic Research.

[28] Gentry, M. and Li, T. (2014). Identification in auctions with selective entry. *Econometrica*, 82(1):315–344.

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

[30] Haile, P. A. (2001). Auctions with resale markets: An application to u.s. forest service timber sales. *American Economic Review*, 91(3):399–427.

[31] Haile, P. A. and Tamer, E. (2003). Inference with an incomplete model of english auctions. *Journal of Political Economy*, 111(1):1–51.

[32] Hortaçsu, A. and McAdams, D. (2010). Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the turkish treasury auction market. *Journal of Political Economy*, 118(5):833–865.

[33] Hörner, J. and Sahuguet, N. (2007). Costly Signalling in Auctions1. *The Review of Economic Studies*, 74(1):173–206.

[34] Ichimura, H. (1991). Semiparametric Least Squares (sls) and Weighted SLS Estimation of Single- Index Models. Papers 264, Minnesota - Center for Economic Research.

[35] Jeziorski, P. (2014). Estimation of cost efficiencies from mergers: application to US radio. *RAND Journal of Economics*, 45(4):816–846.

[36] Jullien, B. and Pavan, A. (2013). Platform Pricing under Dispersed Information. TSE Working Papers 13-429, Toulouse School of Economics (TSE).

[37] Krasnokutskaya, E. (2012). Identification and estimation of auction model with two-dimensional unobserved heterogeneity*. *International Economic Review*, 53(3):659–692.

[38] Krasnokutskaya, E. and Seim, K. (2011). Bid preference programs and participation in highway procurement auctions. *American Economic Review*, 101(6):2653–86.

[39] Levin, D. and Smith, J. L. (1994). Equilibrium in auctions with entry. *The American Economic Review*, 84(3):585–599.

[40] Li, T. and Perrigne, I. (2003). Timber sale auctions with random reserve prices. *The Review of Economics and Statistics*, 85(1):189–200.

[41] Li, T. and Zheng, X. (2009). Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions. *The Review of Economic Studies*, 76(4):1397–1429.

[42] Li, T. and Zheng, X. (2012). Information acquisition and/or bid preparation: A structural analysis of entry and bidding in timber sale auctions. *Journal of Econometrics*, 168(1):29 – 46.

[43] Liu, Q. and Serfes, K. (2013). Price Discrimination in Two-Sided Markets. *Journal of Economics & Management Strategy*, 22(4):768–786.

[44] Matzkin, R. L. (2007). Nonparametric identification. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 6 of *Handbook of Econometrics*, chapter 73. Elsevier.

[45] Roberts, J. W. and Sweeting, A. (2013). When should sellers use auctions? *American Economic Review*, 103(5):1830–61.

[46] Rochet, J.-C. and Tirole, J. (2003). Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, 1(4):990–1029.

[47] Ryan, S., Bajari, P., and Hong, H. (2005). Identification and Estimation of Discrete Games of Complete Information. Computing in Economics and Finance 2005 53, Society for Computational Economics.

[48] Samuelson, W. F. (1985). Competitive bidding with entry costs. *Economics Letters*, 17(1):53 – 57.

[49] Song, M. (2013). Estimating Platform Market Power in Two-Sided Markets with an Application to Magazine Advertising. 2013 meeting papers, Society for Economic Dynamics.

[50] Spence, M. (1976). Product Differentiation and Welfare. *American Economic Review*, 66(2):407–14.

[51] Szerman, D. (2012). *Public Procurement Auctions in Brazil*. London School of Economics and Political Science (University of London).

[52] Tamer, E. (2003a). Incomplete Simultaneous Discrete Response Model with Multiple Equilibria. *Review of Economic Studies*, 70(1):147–165.

[53] Tamer, E. (2003b). Incomplete Simultaneous Discrete Response Model with Multiple Equilibria. *The Review of Economic Studies*, 70(1):147–165.

[54] Tang, X. (2011). Bounds on revenue distributions in counterfactual auctions with reserve prices. *The RAND Journal of Economics*, 42(1):175–203.

[55] Weyl, E. G. (2010). A Price Theory of Multi-sided Platforms. *American Economic Review*, 100(4):1642–72.

[56] Ye, L. (2007). Indicative bidding and a theory of two-stage auctions. *Games and Economic Behavior*, 58(1):181 – 207.