UNIVERSITY OF ECONOMICS IN BRATISLAVA
FACULTY OF NATIONAL ECONOMY

Evidence Number: 101003/I/2024/36122176493196804

ROBUSTNESS OF THE ENTRY'S EFFECT ON THE INTENSITY OF COMPETITION IN PHARMACY RETAIL MARKET

Master Thesis

2024 **Bc. Juraj Šimčisko**

UNIVERSITY OF ECONOMICS IN BRATISLAVA
FACULTY OF NATIONAL ECONOMY

ROBUSTNESS OF THE ENTRY'S EFFECT ON THE INTENSITY OF COMPETITION IN PHARMACY RETAIL MARKET

Master Thesis

Study program: Applied Economics

Field of Study: Economics and management

Department: Department of Economic Policy

Supervisor: Richard Kališ, Dipl. Ing., PhD.

Bratislava 2024 **Bc. Juraj Šimčisko**

“Ehrgeiz ist der Tod des Denkens.”

Ludwig Josef Johann Wittgenstein

Acknowledgements

I would like to express my gratitude to Ing. Richard Kališ, PhD. whose expertise and knowledge were invaluable during this research.

My sincere thanks also go to the Department of Economic Policy for the data used in this thesis.

Together, I must acknowledge my family for their unwavering love and encouragement. Without your support, this project would not have been possible.

ABSTRACT

ŠIMČISKO, Juraj: *Robustness of the entry's effect on the intensity of competition in pharmacy retail market*. – University of Economics in Bratislava. Faculty of national economy; Department of Economic Policy. – Thesis supervisor: Ing. Richard Kališ, PhD. Bratislava: NHF EU, 2024, 73 p.

This thesis investigates the impact of firm entry on market outcomes of over-the-counter and prescribed medicine within the retail pharmacy sector of the Slovak Republic. Employing the entry model developed by Bresnahan and Reiss alongside a reduced form model for observed market-level outcome, we analyse shifts in market dynamics following entry. Leveraging Heckman correction, we compare the findings of both models to draw robust conclusions regarding market efficiency. Our analysis suggests that entry fosters a move towards competitive behaviour, resulting in diminished variable profits. Simultaneously our findings indicate that the Bresnahan and Reiss model can serve as a reliable framework for analysing efficiencies in local markets when data availability is limited. Overall, our results indicate a lack of significant inefficiencies or overconcentration within the Slovak retail pharmacy market once more than four firms entered the market.

Keywords:

competition, pharmacies, entry, robustness, quantity

ABSTRAKT

ŠIMČISKO, Juraj: *Robustnosť vplyvu vstupu na intenzitu konkurencie v maloobchode lekární*. – Ekonomická univerzita v Bratislave. Národohospodárska fakulta; Katedra hospodárskej politiky. – Školiteľ záverečnej práce: Ing. Richard Kališ, PhD. Bratislava: NHF EU, 2024, 73 s.

Táto práca skúma vplyv vstupu firiem na trh v sektore maloobchodných lekární v Slovenskej republike. Využívajúc model vstupu, ktorý vyvinuli Bresnahan a Reiss, spolu s OLS modelom pre pozorované množstvá na úrovni trhu, analyzujeme zmeny v dynamike trhu po vstupe. Využitím Heckmanovej korekcie porovnávame zistenia oboch modelov s cieľom vyvodiť robustné závery týkajúce sa efektívnosti trhu. Z našej analýzy vyplýva, že vstup na trh podporuje prechod ku konkurenčnému správaniu, čo vedie k zníženiu variabilných ziskov. Závery našej práce naznačujú, že model Bresnahan a Reiss môže slúžiť ako spoľahlivý rámec na analýzu efektívnosti na miestnych trhoch, keď je dostupnosť údajov obmedzená. Celkovo naše výsledky naznačujú, že na slovenskom trhu maloobchodných lekární nedochádza k výraznej neefektívnosti alebo nadmernej koncentrácii, keď na trh vstúpili viac ako tri štyri firmy.

Kľúčové slová:
konkurencia, lekárne, vstup, robustnosť, množstvo

Contents

List of Figures

List of Tables

List of Abbreviations

B&R	Model based on the Bresnahan and Reiss (1991) method.
ET	Entry threshold.
ETR	Entry threshold ratio.
MLE	Maximum likelihood estimator.
OTC	Over-the-counter medicine.
PM	Prescribed medicine.

Introduction

Pharmacy retail markets play a pivotal role in healthcare systems, influencing access to essential medications and healthcare services (Smatana et al., 2016). Over the past two decades, the Slovak pharmacy retail market has undergone significant transformations due to multiple waves of deregulation (Kališ, 2023). Understanding the implications of these changes on market dynamics is vital for optimizing resource allocation and enhancing healthcare outcomes.

In this thesis, we analyse population cross-section data on market outcomes, market structures, and demand variables for the pharmaceutical retail market in 2016. Our main objective is to explore the relationship between market structure, demand size, and competition. We aim to investigate how market outcomes in the pharmacy retail sector respond to shifts in market structure.

To enhance the robustness of our findings, we employ two distinct methodological approaches, namely the entry model based on the Bresnahan and Reiss (1991) method, and an Ordinary Least Squares model with dummy variables. Additionally, we draw inferences about the competitiveness of the Slovak retail pharmacy markets in relation to their corresponding market structures.

Our thesis consists of theoretical and empirical sections. In the theoretical section, we provide a theoretical background of the models utilized. Firstly, we present the Bresnahan and Reiss model, followed by a discussion of the theoretical approach employing Ordinary Least Squares with dummy variables.

The empirical section is divided into methodology, results, discussion, and conclusion. The methodology section outlines the methods utilized, including descriptions of limited dependent variable models, ordered least squares method with dummy variables, and theory of regressions with sample selection followed by Heckman correction.

In the results section, we present estimates derived from both the Bresnahan and Reiss model and OLS models, focusing on market outputs for prescribed medicine, over-the-counter medicine, and their respective summaries. Last, we discuss the results, where we deliberate some limitations of our investigation setup, before drawing our conclusions.

Our main findings are that pharmacy retail markets in Slovakia are moving towards the state of perfect competition once entry occurred. This finding indicates absence of concentration in those markets. Additionally, by fifth entry, we do not observe substantial advances to the efficiency of these markets.

With respect to the literature regarding entry models, we have found that the model developed by Bresnahan and Reiss serves as a reliable framework for analysis competition in local markets. To these ends we utilized Ordinary-Least-Squares model with dummy variable to evaluate robustness of estimates from both models. Our analysis proved similar outcomes of both methodologies.

The novelty of our work lies in utilization of Ordinary-Least-Squares as a mean to evaluate robustness of method developed by Bresnahan and Reiss. To the best of the author's knowledge, this method has not been previously applied in this manner.

1 Literature overview and theoretical frameworks

Competition in markets stands at the forefront of economic disciplines, especially in industrial economics, where numerous models have been devised to effectively estimate its levels within specific industries. We position our thesis to the literature of empirical entry models.

For instance, notable studies such as Reiss, P. C., & Spiller, P. T. (1989), which examined the modelling of determinants of competition in airline routes, Berry, S. T. (1992 focused on entry decisions of airlines concerning underlying profitability, and Mazzeo, M. J. (2002), which addressed product-type decisions of firms in oligopoly markets, have contributed significantly to this discourse.

With the escalating significance of healthcare spending, it comes as no surprise that a multitude of studies have emerged, delving into the impact of market structure on competition and economic performance within the healthcare sector (Lábaj et. al., 2018). The seminal empirical investigation into the relationship between market size and market structures in the healthcare sector was pioneered in (Newhouse et al., 1982), focusing on the physician market in the United States.

Subsequently, in 1991, Bresnahan and Reiss, (Bresnahan & Reiss, 1991) published a foundational article on entry literature, exploring the relationship between market structure and competition across various sectors in the US, including the professions of doctors and dentists.

Expanding upon this seminal work, Abraham et al. (2007), utilized discrete factor approximation and Heckman selection to enhance the static entry model proposed by Bresnahan and Reiss. This refinement allowed for the separation of the identification of per-firm changes in costs and the toughness of competition.

Likewise, Schaumans and Verboven (2008) examined the effects of entry and conduct regulation on Belgium pharmacies and physicians, while later, they (Schaumans and Verboven (2011)) proposed a methodology for estimating competition entry effects for differentiated firm products.

The discourse on entry models transcends international boundaries, with numerous domestic studies contributing to this literature. Labaj et al. (2018), for instance, revisited this topic, examining the evolution of market size thresholds across three distinct time periods (1995, 2001, and 2010) within municipality dentist, physician, and pharmacist markets. Their findings shed light on the significant impact of entry into the pharmaceutical sector throughout the transition of the Slovak markets.

Similarly, Kališ (2023) investigated the effects of the liberalization of the Slovak retail pharmacy market. By employing entry models, his results suggest that liberalization led to an increase in the number of pharmacies entering the market, resulting in potentially greater coverage across different market structures.

1.1 Bresnahan and Reiss model

This chapter presents a concise overview of the Bresnahan and Reiss (B&R) model, which forms the cornerstone of our empirical analysis. Here, we introduce the key components and methodologies outlined in their seminal work. Throughout this chapter, our aim is to present the model and its relevance to our thesis by paraphrasing the content from the original paper.

Introduced by Bresnahan and Reiss in 1991, the B&R model offers an empirical framework tailored to assess the impacts of market entry within concentrated markets. Their approach, as detailed in their paper, involves the utilization of an ordered probit model aimed at examining firm entry across 202 relatively compact and geographically isolated markets spanning various professions, including dentistry, medicine, pharmacy, plumbing, and tire retailing. By exploring the effects of entry on these markets, their model facilitates the empirical estimation of competitive behaviour across diverse market compositions.

At the core of their methodology lies the utilization of an ordered probit model, enabling the analysis of how shifts in market demand correspond to alterations in the equilibrium number of firms within these markets. This analysis revolves around the concept of a zero-profit equilibrium level of demand based on the free entry condition, denoted as "entry thresholds" (ET), which signify the minimum market size required to sustain a specific number of firms.

We will employ this model to capture cross-sectional variations in the number of firms within the retail pharmaceutical market of the Slovak Republic in 2016, reflecting the outcomes of entry decisions by these enterprises. Central to this framework is the outcome function of the market, analogous to the market demand function, which can be formalized as:

$$Q = d(\boldsymbol{Z}, P) \times S(\boldsymbol{Y}) \tag{1.1}$$

In this context, $d(\boldsymbol{Z}, P)$ represents the demand function of a "representative consumer", with $S(\boldsymbol{Y})$ denoting the market size, which can be either the number of consumers or another characteristic indicative of market size. Vectors $\mathbf{Y}$ and $\mathbf{Z}$ contain demographic variables that influence market demand, while P signifies price. Vector $\boldsymbol{Z}$ captures the individual demand characteristics within a market, defining the demand characters of markets representative consumer. Conversely, vector $\boldsymbol{Y}$ characterise the market size, typically representing the population size, thereby capturing the outward shifting of the demand curve.

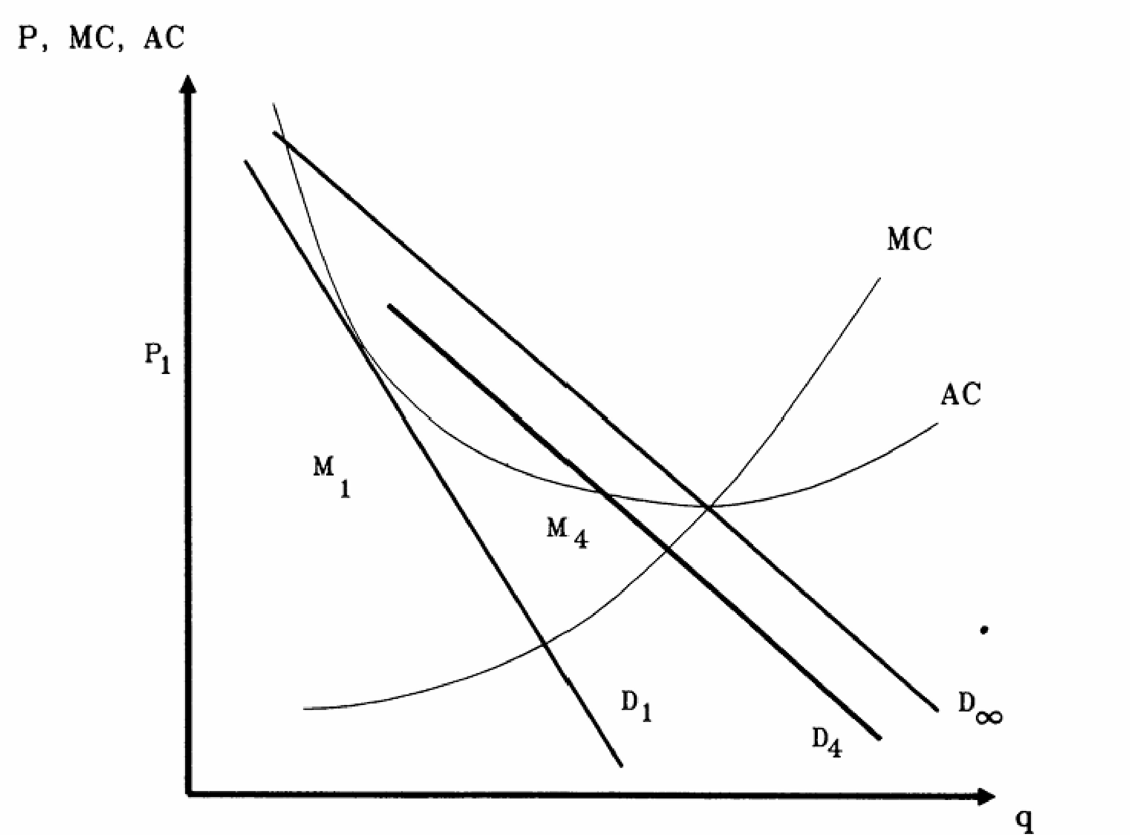
It's important to underscore that those alterations in market size $\boldsymbol{Y}$ lead to proportionate changes in saturated demand Q at any given price P, signifying that price elasticity remains relatively inelastic concerning market size $\boldsymbol{S}$ and only fluctuates in response to the number of entrants N.

On the cost side, the theoretical framework of B&R model incorporates firms fixed costs $F(\boldsymbol{W})$ and marginal costs $MC(q, \boldsymbol{W})$. Where vector $\boldsymbol{W}$ includes exogenous variables influencing costs and q is firm's output. Model assumes average variable cost $AVC(q, \boldsymbol{W})$, that exhibit U-shaped curve. This is consistent with the standard economic theory. Initially higher average variable costs AVC stems from fixed costs, while subsequent increases are driven by rising marginal costs MC.

The objective of the model is to derive insights into the level of competition by establishing a relationship between the number of entrants N and the size of their respective markets S. To examine how N should change with S, we assume markets under the following assumptions regarding market characteristics:

- Homogeneous Products: All products offered within the market are identical.
- Identical Firms: The firms operating within the market are identical in terms of their technologies and capacities.
- Closed, Well-Defined Markets: Markets are considered closed systems with clearly delineated boundaries.
- Saturation of Demand: Changes in the number of firms lead to the saturation of the same demand.

Figure 1: Breakeven firm demand and margins.

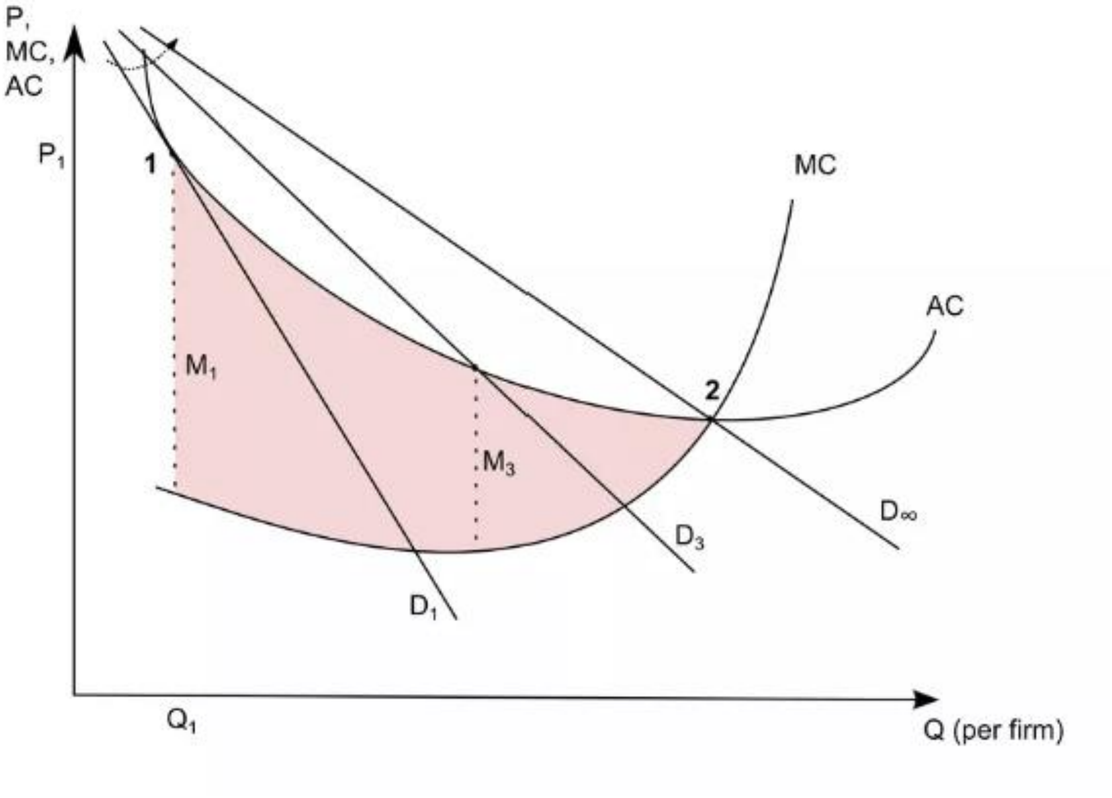


Source: *(Bresnahan and Reiss, 1991)*

Based on the provided assumptions, each firm operates under a long-run cost function (AC), as illustrated in Figure 1. The demand curve labelled (D_1) delineates the minimum level of demand required for a single firm to achieve a break-even point. At this level of demand (S_1), consumers pay a price denoted as P_1. Notably, while a monopolist merely breaks even at this price point, it manages to accumulate a substantial price-cost margin (M_1). This margin represents the disparity between the price paid by consumers and the cost incurred by the monopolist to produce and supply the product.

The model's assumptions suggest that as the market size increases, the anticipated outward rotation of the market demand curve leads to an expansion of the monopolist's profits. Additionally, this growth in market demand enhances the potential post-entry profits for prospective entrants. Eventually, as market demand expands significantly relative to the minimum efficient scale, the price-cost margins of firms tend to converge towards competitive levels. In Figure 2, this convergence is observed when the per-firm demand curve (D_∞) intersects with the minimum point of the average total cost curve.

Figure 2: Firms' margins area.



Source: *(Balmer, 2014)*

To estimate the rate at which oligopoly margins diminish towards zero, our ideal scenario would involve observing how rapidly the breakeven price-cost margins $M_N = P_N - MC(q_N)$ decrease as the number of firms (N) increases (red area in Figure 2.) with each entry. It's crucial to note that an entrant will only participate in the market if the margins satisfy the condition of at least zero economic profits. This condition can be formalized as $M_N = P_N - MC \geq 0$. Additionally, with each new entrant, these margins are reduced, implying that $M_N > M_{N+1}$.

In the absence of margin data, reliance on entry thresholds serves as a methodological approach to infer about margins. To elucidate the implications of entry thresholds on the effects of entry, we start by comparing the monopoly and competitive entry thresholds. A monopolist attains zero economic profits when its fixed costs are offset by variable profits, effectively reflecting the condition of marginal zero economic profits. This condition can be formalized as follows:

$$\Pi_1(S_1) = [P_1 - AVC(q_1, \boldsymbol{W})]d(\boldsymbol{Z}, P_1)S_1 - F = V^* - F = 0 \qquad (1.2)$$

Where the entry threshold for monopolists (S_1) represents the minimum market size required to sustain the monopolist's unobserved fixed costs (F) through variable profits (V).Note that variable profits V^* are per customer profits $[P_1 - AVC(q_1, \boldsymbol{W})]d(\boldsymbol{Z}, P_1)$ multiplied by market size or number of customers and those are calculated as per unit profit $[P_1 - AVC(q_1, \boldsymbol{W})]$ multiplied by characteristics of representative consumer’s demand $d(\boldsymbol{Z}, P_1)$. We can rearrange equation (1.2) to derive its corresponding equation:

$$S_1 = \frac{F}{[P_1 - AVC(q_1, \boldsymbol{W})]d(\boldsymbol{Z}, P_1)} \qquad (1.3)$$

Expression (1.3) offers valuable insights into the market's characteristics. Firstly, firms operating in markets with elevated fixed costs necessitate a larger market size for entry. Secondly, a larger variable probit correlates with a diminished entry threshold. Lastly, heightened representative consumer demand corresponds to a reduced entry threshold.

The competitive equivalent of the monopoly entry threshold is the per-firm entry threshold (s_∞)[1] defined as $s_\infty = \lim_{N\to\infty} S_N/N$. The ratio of entry thresholds, s_∞/s_1, measures the decrease in variable profits per customer between a monopoly and a competitive market. This scale free metric ranges from marginal costs upwards and increases as the monopolist's

---

[1] Note the distinction between s and S, where s refers to per-firm entry threshold.

demand curve becomes steeper. (see Figure 2). In other words, the more profitable a monopolist is, the higher s becomes.

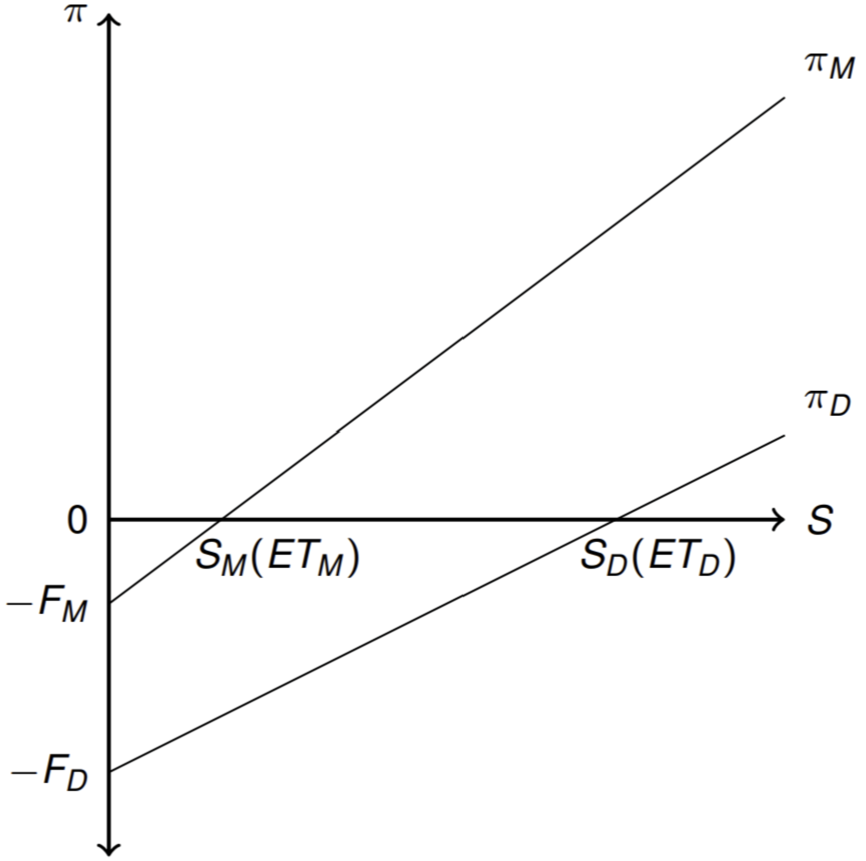
1.1.1 From monopoly to oligopoly

The objective of Bresnahan and Reiss's model is to derive insights into alterations in competition concerning shifts in market structures. An analysis based on Section 1.1 indicates that we can utilize ratios of oligopoly entry thresholds to infer changes in oligopoly margins. In this context, the work of Bresnahan and Reiss furnishes us with a pivotal example that elucidates the operational mechanics of this phenomenon. We hereby paraphrase this illustrative example.

Suppose that it takes 2,000 customers to sustain a monopolist ($S_1 = 2\ 000$) and the market transitions to perfect competition when each firm serves 4,000 customers ($S_\infty = 4\ 000$). These two entry thresholds encompass the spectrum of oligopoly entry thresholds we anticipate observing. For instance, if the fourth entrant anticipates competing in a perfectly competitive environment, we should observe $S_4 = 4 \times 4\ 000 = 16\ 00$ customers served, or $S_\infty / S_4 = 1$. This ratio indicates that quadropolists earn the same variable profits per customer as competitive firms. Alternatively, if the fourth entrant operates within a cartel and enters the market when covering its fixed costs at the monopoly price, it enters when the market has $4 \times 2\ 000 = 8\ 000$ consumers, resulting in S_4/S_∞=2. Extending this rationale to degrees of post-entry competition between cartels and perfect competition, we would typically expect to observe per-firm entry thresholds ranging between 2,000 and 4,000 customers. For instance, observing $S_4 = 3\ 810$ consumers served would often lead to the conclusion that the market is nearly competitive. In such a scenario, the ratio S_4/S_∞ equals 1.05, indicating that a quadropolist serves approximately 5% fewer customers than a competitive firm. Or the difference between the state of perfect competition is $\left[s_\infty = \lim_{N\to\infty} S_N/N\right] - [S_4/S_\infty] = 1 - 1.05 \ \Rightarrow\ \Delta 5\%$.

Note that this illustrative example, entirely taken from Bresnahan and Reiss (1991), does not account, as well as our analysis, for changes in fixed costs. According to standard economic theory, we anticipate increasing entry barriers with each additional entrant. The relationship between market size and rising fixed costs can be depicted in Figure 3. The monopolist's profit function, π_M , ascends from its fixed costs F_M, effectively satisfying the zero-economic profit condition at the monopolist's market size S_M. However, as the second entrant confronts higher variable and fixed costs F_D, its entry threshold,ET_D and profit's slope π_M is not proportional to the monopolist's profits and entry threshold ET_M.

Figure 3: Relationship of market size and profits (entry barriers assumed)



Source: *(Kališ, 2023)*

We can formalize the preceding information concerning oligopolistic entry thresholds and develop the function of the profits for the Nth entrant in expression (1.4). By applying postulates about markets, lets recall, that the model assumes identical firms with homogenous products operating within shared defined market and market demand can be equally split between each entrant.

$$\Pi_N = [P_N - AVC(q_N, \boldsymbol{W}) - b_N] d(\boldsymbol{Z}, P_N) \frac{S}{N} - F_N - B_N \tag{1.4}$$

In original paper the constant $b_N \geq 0$ and $B_N \geq 0$ in equations (1.4) to (1.7) provide a solution to the issue of increasing entry barriers for later entrants. While we do not utilize them in the empirical section of our thesis, this aspect of Bresnahan and Reiss's model is noteworthy in the theoretical part.

Similarly, as in expression (1.3), the breakeven condition $\Pi(S_N) = 0$ defines breakeven level of demand for Nth entrant, which we refer to as the per-firm entry threshold. Formally:

$$s_N = ET_N = \frac{S_N}{N} = \frac{F_N + B_N}{(P_N - AVC_N - b_N) d_N} \tag{1.5}$$

Expression (1.5) bears a resemblance to expression (1.3), the constants B_N and b_N serve to account for disparities between the entrant's variable and fixed costs. Similarly, we can draw inferences about the characteristics of the per-firm entry threshold, ET_N, based on equation (1.5). Holding production and entry costs constant, we note that s_N decreases with an increase in variable profits and a decrease in fixed costs.

1.1.2 Entry threshold ratios

Following previous chapters, we can demonstrate that if firms possess identical costs and if entry does not alter competitive behaviour, then $S_{N+1}/S_N = 1$. Consequently, deviations of successive entry threshold ratios from 1 indicate changes in competitive conduct as the number of firms increases. Indeed, this statistic does not directly measure the level of competition. Instead, it evaluates how this competitive levels changes with the market structure (number of firms N). The expectation is that in a cartel, the Entry Thresholds

(s_N) would remain constant as the number of firms changes. Conversely, in competitive markets, we anticipate convergence to 1.0 as the number of firms increases.

As entry thresholds are calculated using variables $\boldsymbol{Z}$, $\boldsymbol{Y}$ and $\boldsymbol{W}$ which vary across different markets, employing entry thresholds to draw conclusions about diverse markets proves impractical. Therefore, B&R model introduces the concept of the Entry Threshold Ratio (ETR), formally expressed as:

$$ETR = \frac{S_N}{N} \Big/ \frac{S_{N+1}}{N+1} = \frac{ET_{N+1}}{ET_N} = \frac{V_N}{V_{N+1}} \times \frac{F_{N+1} + B_{N+1}}{F_N + B_N} \qquad (1.6)$$

In equation (1.6) V_N stands for Nth entrant's breakeven variable profits and $F_N + B_N$ similarly information about Nth fixed costs. The ETR measures the rate at which fixed costs or variable profits decrease with entry, illustrating how the level of competition evolves with the number of firms. In the case of perfect competition, the ETR would maintain a value of 1.00 across all market structures, signifying that every potential market size is saturated by new entrants without incurring additional costs. However, such a scenario is unrealistic for real-world markets, as every firm possesses fixed costs that must be covered by market size to fulfil the zero-profit condition. In competitive markets the ETR should approach 1.00, indicating a relative state of perfect competition with each additional entrant. Conversely, markets where this pattern is disrupted exhibit signs of overconcentration, potentially leading to deadweight losses.

Finally, by relaxing the condition of increasing entrants' barriers, we can eliminate the constants B_N and b_N. We then rearrange expression (1.6) to its final form, as utilized in our empirical analysis, as follows:

$$ETR = \frac{S_N}{N} \Big/ \frac{S_{N+1}}{N+1} = \frac{ET_{N+1}}{ET_N} = \frac{F_{N+1}}{F_N} \frac{(P_N - AVC_N) \times d_N}{(P_{N+1} - AVC_{N+1}) \times d_{N+1}} \qquad (1.7)$$

1.2 Ordinary least squares models

To measure the robustness of the B&R approach, Abraham et al. (2007) proposed an additional method. In their methodology, Abraham, Gaynor, and Vogt establish a link between the Bresnahan and Reiss model and an Ordinary Least Squares (OLS) model for realized quantities. Through the discrete factor approximation, this approach allows for the separation of the underlying factors in the Bresnahan and Reiss model, including the change in per capita demand, average variable profits, and fixed costs with entry. Unfortunately, their primary objective was accomplished using Tomas Mroz's discrete factor approximation, which the author was unable to employ in this proposition.

In this thesis, our objective is to employ this concept in a novel manner. As outlined in Bresnahan and Reiss's example, the entry threshold ratios are construed as indicators of how swiftly markets transition to competitive performance. We can refine this definition to interpret them as measures of how effectively the market demand is saturated in response to changes in market structures. From this perspective, it becomes clear that simple Ordinary Least Squares (OLS) with dummy variables can reflect the Entry Threshold Ratios (ETRs) calculated in the Bresnahan and Reiss model.

To show how, we should restate the demand function proposed by B&R (1.1). Where we see that the demand function is represented as a representative consumer demand multiplied by market size. In this model, however, our aim is not to observe nor estimate the entry thresholds as a minimal market size necessary to support market structure with N firms. Instead, we use demand function to evaluate how demand saturation changes with respect to changes in market structures. To articulate this concept differently, in the B&R model, our objective is to estimate the minimum market size, known as the Entry Threshold (ET). Conversely, in the OLS model, the focus shifts to directly observing changes in market outcomes within different market structures, relative to the demand of the representative customer and the market size. Under the premise of competitive behaviour, it is logical to anticipate a proportion between increasing market saturation and the number of entrants. As additional firms enter the market, the overall saturation of demand is expected to rise proportionally. By examining the degree of proportionality between changes in market outcomes and variations in market structure, holding other demand variable same, we can make inferences about the level of competition within the market.

To illustrate this concept, let's consider a scenario akin to that proposed by Bresnahan and Reiss (B&R). Suppose we have a monopolist saturating market demand at quantity level q_M. In a hypothetical scenario of perfect competition, a doubling of market size with the same demand from the representative consumer would lead to the entry of a new firm. This new entrant would then saturate the demand at an equal level, denoted as $q_D = q_M$.

This straightforward reasoning mirrors the concept of Entry Threshold Ratios (ETR) proposed by B&R. By comparing the ratio of these quantities, $q_{N+1} \,/\, q_N$, we can infer insights into the changing competitiveness of markets with each successive entry. A value exceeding 1.0 indicates that entry has resulted in a level of market demand saturation surpassing that achieved by the monopolist, suggesting inefficiencies within the market. Conversely, a value converging towards 1.0 suggests a gradual convergence towards competitive conduct, where market demand saturation aligns closely with competitive efficiency.

It's important to note that this reasoning is contingent upon B&R assumptions, including constant costs, zero entry barriers, and assumed market characteristic. However, despite fundamental methodological distinctions, the primary practical difference lies in the data requirements. While the B&R model maintains simplicity in data requirements, our OLS model necessitates data on realized quantities, which can be empirically difficult to acquire.

To formalize this theoretical approach, lets rearrange original demand function (1.1). Unlike previously, we observe the realized quantities Q_N for each market structure N. Our point of interest lies in the rate of change in realized quantities with a change in market structure. Indeed, the change in realized quantities between different market structures can be compared with individual entry thresholds for the respective market structures.

$$Q_{N+1} - Q_N = \Delta Q_{N+1} \sim ET_{N+1} \tag{1.8}$$

Similarly, we aim to establish a scale-free measurement for these changes by calculating their ratios. This yields another form of Entry Threshold Ratio (ETR), whose interpretation in this alternative methodological setup mirrors that proposed by Bresnahan and Reiss (B&R).

$$\frac{\Delta Q_N}{\Delta Q_{N-1}} \sim ETR = \frac{ET_N}{ET_{N-1}} \tag{1.9}$$

After estimating these two methodologies, we utilize both sets of estimates as a robustness check for the estimations derived from both models. This comparative analysis functions as a validation mechanism, enabling us to assess the robustness of the estimations obtained through both approaches.

To avoid confusion, this is a theoretical conceptualization of OLS model, entirely formulated by author. For a comprehensive explanation, please refer to Chapter (3.2.5) and Equation (3.27), which clarify the econometrics of this theoretical framework.

2 Aim of the thesis

The aim of this thesis is to enhance the precision of estimating the impact of market entry by incorporating data on realized quantities in the markets. Our investigation focuses on the cross-sectional data of the Slovak retail pharmacy market in 2016. The primary objective can be segmented into the following steps:

In the first step, we employ the Bresnahan and Reiss model described in (Bresnahan and Reiss, 1991) to estimate the respective entry thresholds. This is accomplished through ordered probit model, where we estimate the entry threshold ratios as outlined in the theoretical framework.

In the second step, we estimate dummy variables for market structures using standard Ordinary Least Squares (OLS) with market structure dummies (similar to (Abraham et. al., 2007)) for over-the-counter medication, prescribed medication, and their aggregate quantities. To address an underlying selection bias in this model, specifically incidental truncation, we employ Heckman correction.

In the third and final step, we compare these estimates and draw inferences about the nature of competition with respect to market structures.

The research question is: *how do market outcomes in the pharmacy retail market change with respect to the changes in market structure?*

Addressing the research question through the utilization of two distinct methodologies not only enables an evaluation of competition in the Slovak retail pharmacy market but also facilitates a comparative analysis of these two assessments. This comparative approach acts as a robustness check for the estimations, thereby evaluating the reliability and validity of these models.

3 Methodology

The aim of this chapter is to describe the methodology and econometric aspects of the models employed. This section is divided into two parts. The first part provides a concise overview of our dataset and its sources. The second part is dedicated to presenting the econometric methods used to achieve the objectives outlined in the previous section.

We commence this part by delineating limited dependent variable models, focusing primarily on probit models. Subsequently, in the second part, we expound upon the methodology behind ordered probit models. In the third part, we describe the methodology of the Bresnahan and Reiss model within the context of ordered probit models. The fourth part provides a brief introduction to ordinary least square models with dummy variables. In the last two parts, we illuminate the selection bias present in our methodology and discuss Heckman correction to address this issue.

3.1 Data description

In the empirical part of this thesis, we analyse data concerning medication sales in the Slovak Republic during 2016. Specifically, we examine the quantities of over the counter and prescribed medications sold, as well as their corresponding market summaries, across administratively defined regions (towns). These data were sourced from the Ministry of Healthcare of the Slovak Republic. Additionally, we gathered information on the number of pharmacies operating within these towns during the same period.

Furthermore, to estimate market demand, we incorporated data on various socio-economic factors including the number of households, average wages, demographic distribution (containing the population in different age groups such as young, old, and productive people), and the proportion of unemployed individuals. These socio-economic indicators were collected from the Statistical Office of the Slovak Republic for all towns within the country during 2016.

For estimating fixed costs, we acquired data on pharmacist wages from the Ministry of Healthcare of the Slovak Republic.

3.2 Statistical methods and econometric specifications

The empirical part of our thesis employs the Bresnahan and Reiss model to estimate the impact of entry on competition intensity within pharmacy retail markets. Additionally, we utilize a reduced form Ordinary Least Squares (OLS) model to assess the robustness of our estimations. As outlined in the theoretical section, we apply an ordered probit model for the Bresnahan and Reiss model, and for the OLS model, we utilize a simple OLS model with dummy variables representing respective market structures. However, our setup exhibits selection bias in the OLS analysis. To address this issue, we employ Heckman correction to obtain unbiased estimates for the OLS method.

This methodology section begins with a description of limited dependent variable models, followed by an explanation of the ordered probit model. Subsequently, we detail the econometric specification of the Bresnahan and Reiss model. Then, we proceed to describe ordinary least squares models with dummy variables. Finally, we discuss selection bias and the implementation of Heckman correction to resolve incidental truncation in the OLS method.

Most of the statistical methods described in this section closely resemble the content found in Wooldridge's "Introductory Econometrics: A Modern Approach" from 2012 as this was our main source regarding econometrics theory. (Wooldridge, 2012)

3.2.1 Limited dependent variable models

Binary variables are a typical example of limited dependent variables. Their limitation arises from the restriction imposed on the range of their values, which is strictly between two values, typically zero and one. Similarly , limited dependent variables take on a relatively small number of usually positive integer values. A prime example is the number of firms in a market. Since their value is strictly nonnegative and usually does not exceed large values, treating them as continuous variables using linear probability proves inefficient. The binary variables in this case can be a probability of presence of a firm on a market. To address this issue, we introduce the binary response model.

$$P(y=1|\boldsymbol{x}) = P(y=1|x_1, x_2, \dots, x_k) \tag{3.1}$$

This model estimates the probability of a response, for given x where $\boldsymbol{x}$ represents a set of explanatory variables. In our case, the decision of a firm to enter ($y = 1$) or leave ($y = 0$) a market is based on several factors denoted by $\boldsymbol{x}$, such as market size and demand characteristics. To avoid linearity in the vector $\boldsymbol{\beta}$, we employ the binary response model as follows:

$$P(y=1|\boldsymbol{x}) = G(\beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k) = G(\beta_0 + \boldsymbol{x\beta}) \tag{3.2}$$

Where G stands for function taking on values strictly between zero and one: $0 < G(z) < 1$ for all real numbers z. This so-called distribution function, ($G(z)$), ensures the range of response probabilities strictly within this interval.

For the distribution function G, various nonlinear functions have been proposed. Among these, we employ the probit model in all our analyses. In the probit model, the distribution function is the standard normal cumulative distribution function (CDF), derived from the density function of the normal distribution (hereafter referred to as NDF). The NDF is one of the most used distributions in econometrics and general statistics. It lies at the cornerstone of the central limit theorem, and statisticians often assume that random variables across a population follow a normal distribution, particularly when the random variable is continuous. Mathematically, the probability distribution function (PDF) can be formalised as:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]}, -\infty < x < \infty \tag{3.3}$$

In traditional notation for population $\mu = E(X)$ donates mean value of X and $\sigma^2 = Var(X)$ donates variance of X. We express, that X which has normal distribution with expected value μ and variance σ^2 as $X \sim Normal(\mu, \sigma^2)$. Additionally, normal distribution is symmetrical about μ. Special case of normal distribution is standard normal distribution, which occurs when $\mu = 0$ and $\sigma^2 = 1$. The PDF of a standard normal random variable z is donated as $\phi(z)$ formally as:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\left(\frac{-z^2}{2}\right)} \tag{3.4}$$

The cumulative function of PDF usually donated as $\Phi(z)$ is area under ϕ. Formally $\Phi(z) = P(Z \leq z)$. Given the properties of CDF, we can employ it to compute problems regarding probabilities of event involving standard normal random variables as described in probit models. Notably, the expression $P(Z > z) = 1 - \Phi(z)$ holds significance.

Recall, that in probit model distribution function $G(z)$ is $\Phi(z)$ ensuring, that value x_i lies strictly between zero and one. $G(z)$ increases most quickly at $z = 0, G(z) \rightarrow 0$ as $z \rightarrow -\infty$, and $G(z) \rightarrow 1$ as $z \rightarrow \infty$.

Important property of binary response models is their ability to be derived from latent variable models. In models, where $y^*$is latent (unobserved) variable, such as:

$$y^* = \beta_0 + \boldsymbol{x\beta} + e, y = 1[y^* > 0] \tag{3.5}$$

Where y is one if $y^* > 0$ and y is zero if $y^* \leq 0$, based on indicator function of $1[\cdot]$ The crucial point here is that this model assumes, that e is independent of $\boldsymbol{x}$ and that e has standard normal distribution. As we follow the general assumption about the (symmetrical) normal distribution of e we favour probit model in all our models, rather than logit model. We stress, that unlike in liner models, the interpretation of estimates $\boldsymbol{\beta}$ is very complicated given the nonlinear nature of distribution function G. From expression (3.5) and assumptions about cdf we can derive response probability as:

$$\begin{aligned} P(y = 1|\boldsymbol{x}) = P(y^* > 0|\boldsymbol{x}) &= P(e > -(\beta_0 + \boldsymbol{x\beta}|\boldsymbol{x}) \\ &= 1 - G[-(\beta_0 + \boldsymbol{x\beta})] = G(\beta_0 + \boldsymbol{x\beta}) \end{aligned} \tag{3.6}$$

3.2.2 *Maximum likelihood for limited dependent variable models*

To estimate response model's nonlinear distribution of $E(y|\boldsymbol{x})$, we employ maximum likelihood estimation (hereafter MLE). The goal of MLE, unconditional distribution estimator, is to find optimal way to fit a distribution of data. (Wooldridge, 2012) This technique erose from the fact, that density of y_i is conational on given $\boldsymbol{X}_i$. In case of probit model, we can write that:

$$f(y|\boldsymbol{x}_i;\boldsymbol{\beta}) = [G(\boldsymbol{x}_i\boldsymbol{\beta})]^y[1 - G[\boldsymbol{x}_i\boldsymbol{\beta})]^{1-y}, y = 0; 1 \tag{3.7}$$

From expression (3.7) we see that the response y taking values strictly between zero and one has density as:

$$y\begin{cases} = 1 \Rightarrow \quad G(\boldsymbol{x}_i\boldsymbol{\beta}) \\ = 0 \Rightarrow 1 - G(\boldsymbol{x}_i\boldsymbol{\beta}) \end{cases} \tag{3.8}$$

From expression (3.8) we can derive log-likelihood function for observation i by taking logarithmic form of it as:

$$\ell_i(\boldsymbol{\beta}) = y_i \log[G(\boldsymbol{x}_i\boldsymbol{\beta})] + (1 - y_i) \log[1 - G(\boldsymbol{x}_i\boldsymbol{\beta})] \tag{3.9}$$

For n sample size the log-likelihood is the summarization across all observations $\ell_i(\boldsymbol{\beta})$.For probit model, where G is cdf, we call estimated $\widehat{\boldsymbol{\beta}}$ probit estimators.

$$\mathcal{L}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\beta}) \tag{3.10}$$

3.2.3 Ordered probit model

As explained in the previous section, the use of OLS regression is ineffective for measuring ordered responses. Therefore, we introduced the probit model to estimate binary responses. We can extend these models to account for multiple ordered responses. This model still deals with limited dependent variables, as their range is constrained to integers, and they only take on a small set of values. We rearrange expression (3.6) as $y^* = \boldsymbol{X}'\boldsymbol{\beta} + e$. In this case, y_i^* represents the observed ordinal variable, which takes on values 0 to m according to scheme: ((Wooldridge, 2012), (Greene, 2002) and (Jackman, 2000))

$$y_i = j \ \Leftrightarrow \mu_{j-1} < y_i^* \leq \mu_j \tag{3.11}$$

Where $j = 0, .., m$, $\mu_{-1} = -\infty$ and $\mu_m = \infty$. The point of interest is how changes in the predictor affect probability of observing ordinal outcome. We can formalize it as:

$$P(y_i = 0)\begin{cases} P(\mu_{-1} < y_i^* \leq \mu_0) \\ P(-\infty < y_i^* \leq \mu_0) \\ \quad P(y_i^* \leq \mu_0) \end{cases} \tag{3.12}$$

By substitution from (3.6) we can rewrite the probability of $y_i = 0$ as follows. Note that we assume $e_i \sim N(0,1), \forall i = 1, \dots, N$.

$$P(y_i = 0)\begin{cases} P(\boldsymbol{X_i}\boldsymbol{\beta} + e_i \ \leq \mu_O) \\ P(e_i \leq \mu_0 - \boldsymbol{x}_i\,\boldsymbol{\beta}) \\ \quad \Phi(\mu_0 - \boldsymbol{x}_i\,\boldsymbol{\beta}) \end{cases} \tag{3.13}$$

Similarly, we can expand this logic to $y_i = 1$ as:

$$P(y_i = 1)\begin{cases} P(\mu_0 < y_i^* \leq \mu_1) \\ P(\mu_0 < \boldsymbol{X_i\beta} + e_i \leq \mu_1) \\ P(\mu_0 - \boldsymbol{X_i\beta} < e_i \leq \mu_1 - \boldsymbol{X_i\beta}) \\ \Phi(\mu_1 - \boldsymbol{x}_i\boldsymbol{\beta}) - \Phi(\mu_0 - \boldsymbol{x}_i\boldsymbol{\beta}) \end{cases} \tag{3.14}$$

Analogy to these responses is general model as follows:

$$P(y_I = j) = \Phi\left(\mu_j - \boldsymbol{X}_i\boldsymbol{\beta}\right) - \Phi(\mu_{j-1} - \boldsymbol{X}_j\boldsymbol{\beta}) \tag{3.15}$$

Except for $j = m$, where:

$$P(y_i = m) = 1 - \Phi(\mu_{m-1} - \boldsymbol{x_i\beta}) \tag{3.16}$$

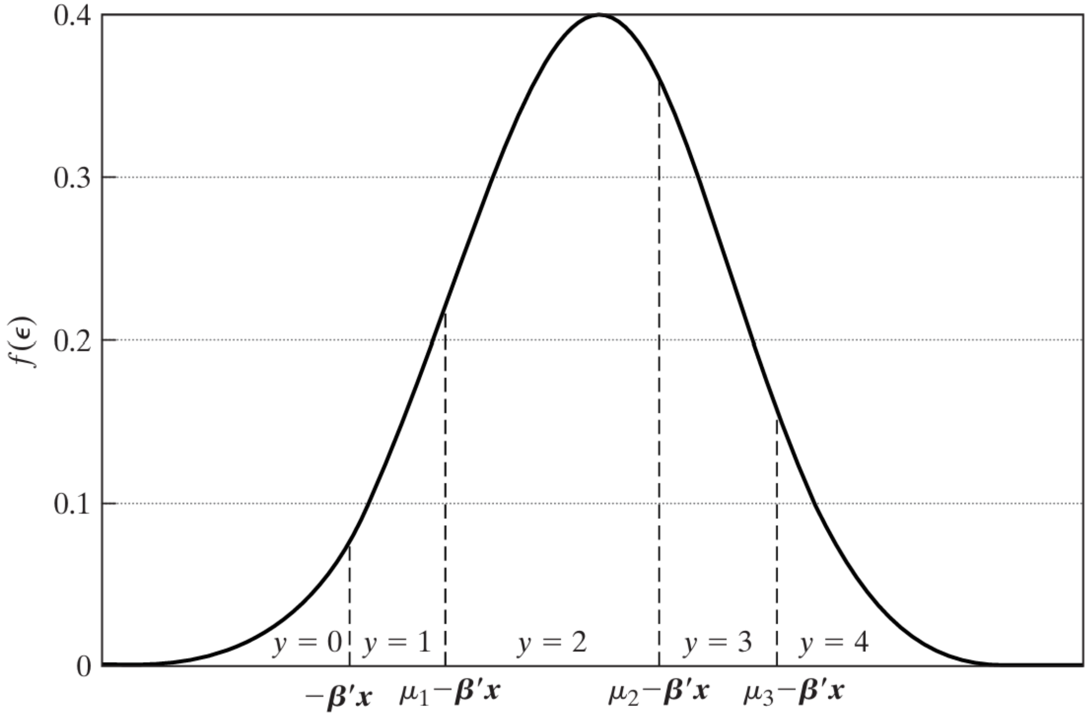
We can effectively simplify expressions (3.13) to (3.16) as:

$$y = \begin{cases} 0 & if & y^* \leq 0, \\ 1 & if & 0 < y^* \leq \mu_1, \\ 2 & if & \mu_1 < y^* \leq \mu_2 \\ \vdots & \vdots & \vdots \\ m & if & \mu_{m-1} < y^* \end{cases} \tag{3.17}$$

For visualization purposes we present expression (3.17) as a figure 4 of probabilities in ordered probit model. The MLE estimator like one for probit model captured in expressions (3.9) and (3.10) is:

$$\begin{gathered} \log \mathcal{L} = \sum\nolimits_{i=0}^{N} \sum\nolimits_{j=0}^{m} Z_{ij} \log\left(\Phi_{ij} - \Phi_{\mathrm{i,j-1}}\right), where \\ \Phi_{ij} = \Phi\left(\mu_j - \boldsymbol{x}_i\boldsymbol{\beta}\right) \ and \ \Phi_{i,j-1} = \Phi(\mu_{j-1} - x_{\boldsymbol{i}}\boldsymbol{\beta}) \end{gathered} \tag{3.18}$$

Figure 4: Probabilities in the Ordered Probit Model



Source: *Greene (2002)*

3.2.4 Econometric specification of B&R model

In this section, we aim to describe the econometrics of the Bresnahan and Reiss model. As outlined in the theoretical part, our objective is to estimate entry thresholds to draw inferences about how competition changes with entry. To achieve this, we utilize the zero-economic profit condition outlined in expression (1.4). Since our model assumes constant costs, we set the constants B_N and b_N equal to 0. We rewrite this expression as a discrete choice model of entrant's long-run profits (3.17), where we expect the number of entrants N when $\boldsymbol{\Pi}_N \geq 0$ and $\boldsymbol{\Pi}_{N+1} < 0$ as follows: ((Bresnahan and Reiss, 1991), (Lábaj et. al. 2018))

$$\Pi_N = S(\boldsymbol{Y}, \lambda) V_N(\boldsymbol{Z}, \boldsymbol{W}, \alpha, \beta) - F_N(\boldsymbol{W}, \gamma) + \varepsilon = 0 \tag{3.19}$$

Where α, β, λ and γ are profit function parameters, $\boldsymbol{Y}$ represents market size, $\boldsymbol{Z}$ and $\boldsymbol{W}$ donates per capita demand and cost shifters. ε represents the market shock, i.e. unobserved profits. This model utilizes orders probit model and can be rewrite as:

$$\widehat{\Pi}_N = \boldsymbol{X}'\beta - \varepsilon \tag{3.20}$$

From an econometric standpoint, this equation represents an ordinal response model employing the probit distribution function, given the similarity in assumptions regarding market shocks and the error term in the probit case. Here, we estimate latent profits ($\widehat{\Pi}_N$) based on the observed number of firms (e.g., market structure). We can further rearrange expression (3.20) into a simplified ordered probit model, as shown in expression (3.17). It's worth noting the change in the notation of the probability threshold from μ_j to δ_N, a typical notation in models concerning the utilization of response models in industrial economics.

$$N = \begin{cases} 0 & if & \widehat{\Pi}_0 \leq 0, \\ 1 & if & 0 < \widehat{\Pi}_1 \leq \delta_1, \\ 2 & if & \delta_1 < \widehat{\Pi}_2 \leq \delta_2 \\ \vdots & \vdots & \vdots \\ m & if & \delta_{m-1} < \widehat{\Pi}_{m-1} \end{cases} \tag{3.21}$$

Where N is the number of firms on market and $\widehat{\Pi}_N$ are latent firm's profits on corresponding market structure. For effective calculations, we can rearrange expression (3.21) in a single equation capturing the main idea of B&R model as follows:

$$\widehat{\Pi}_N = \ln \lambda(S_N) + \boldsymbol{X}'\beta - \delta_N = 0 \tag{3.22}$$

From this expression, we observe that the latent profit of a firm operating in a market with N firms satisfy the zero-economic profit condition. It is calculated based on the market size parameter λ and the parameters $\boldsymbol{\beta}$ capture demand $\boldsymbol{Z}$ and cost $\boldsymbol{W}$ shifters from expression (3.19). The parameter δ_N captures the probability threshold ratios obtained from the ordered probit model. The logarithm of the parameter λ signifies that this model uses a log-log form.

Respective entry threshold, described in (1.5), are then calculated by rearranging expression (3.22) as follows. Note that our model does not count on entry barriers.

$$S_N = e^{\left(\frac{-\boldsymbol{X}'\beta+\delta_N}{\lambda}\right)} \qquad (3.23)$$

The objective of the Bresnahan and Reiss model is to estimate scale-free measurements known as entry threshold ratios. By substituting (1.7) into (3.23), we can estimate them by leveraging the previous results obtained from the ordered probit model.

$$ETR = \frac{S_N}{N}\Big/\frac{S_{N+1}}{N+1} = \frac{e^{\left(\frac{-\boldsymbol{X}'\beta+\delta_N}{\lambda}\right)}}{e^{\left(\frac{-\boldsymbol{X}'\beta+\delta_{N+1}}{\lambda}\right)}} \qquad (3.24)$$

3.2.5 OLS model

The Bresnahan and Reiss model introduces Entry Threshold Ratios (ETRs) to estimate changes in market outcomes concerning the number of firms in the market. However, relying solely on these estimates based on a vector of market characteristics may not provide sufficient persuasion. The estimation heavily depends on the comprehensive choice of variables in expressions (3.19) and (1.1) respectively, which might not capture all relevant market characteristics. Moreover, methodologically, B&R model is based on strong assumptions, namely the assumption of normality and independence of e. To address this, we aim to verify the robustness of B&R estimates by comparing them with standard Ordinary Least Squares (OLS) estimations.

In this method, we regress the market outcome, described as the actual quantity realized in the markets, on a vector of market characteristics (the same $\boldsymbol{X}'\boldsymbol{\beta}$ as in the B&R method) and dummies for market structure, $\boldsymbol{\delta_N}$. Conclusions are drawn based on the comparison of estimations of these dummies. The estimator of the dummy variable can be interpreted as the change in market outcomes relative to the market structure (i.e., the number of firms in the market N). According to expectations based on theory, this estimator tends to increase in markets where no concentration is expected, and the relative change slows with each firm's entry. Eventually achieving competitive conduct with a value their ratio equal to 1, where no difference between respective entry thresholds is observed. This behaviour of the dummy estimator δ_N mirrors the ETR described in the B&R model if the change in fixed and variable costs is not significant (e.g., assumption about non-existent entry barriers).

However, it's crucial to note that the method and estimation process differ. In the B&R model, Maximum Likelihood Estimation (MLE) is employed, while OLS regression is utilized in this method. The dependent variable in the B&R model is the vector N, representing the number of firms, whereas in OLS regression, it is the vector of market outcomes Q. The only common parameters here are the dummies for market structures and the vector of explanatory variables $\boldsymbol{X'\beta}$, which can be perceived as mere control variables (for estimated slope). As the point of interest in the OLS model is not solely on the actual estimations for dummy variables, but rather on analysing the changes of outcomes with respect to variations in market structures. We should be able to obtain similar estimations for the change of outcomes, even though it would be hard to compare them with estimations for B&R and more importantly, we would encounter situation with omitted variables. For this reason, we use the same vector $\boldsymbol{X'\beta}$. We can econometrically formalize this simple OLS model with dummy variables as follows:

$$y = \boldsymbol{X'\beta} + \sum_{i=1}^{N} \delta_N D_N + \varepsilon \qquad (3.25)$$

In this scenario, y represents observed realized quantities Q. The vector of variables $\boldsymbol{X}$ is identical to the variables from expression (3.19), although these variables serve as control variables and their estimation is not our primary focus.

Importantly, the dummy D_N is an observed dummy for respective market structures. The value of D_N is identical to the number of firms operating in the market. We observe markets with a monopolist $D_1 = 1$, as $N = 1$ to mostly penetrated markets $D_N = N$, when $N = \max N$. Thus, we can formalize these dummies as $\sum_{i=1}^{N} \delta_N D_N$. The dummy estimates $\boldsymbol{\delta_N}$ form a vector of estimated changes with respect to market structure. As described, the dummy D_N serves here as a quasi-qualitative factor. Even though it possesses perfect quantitative properties, we treat this variable as if it was categorical variable for a type of market structure.

To illustrate how our model operates, let's consider a simple example as described in expression (3.25) with markets having just one or two firms. Here, y denotes quantity, and D_N represents a dummy for market structure, where 0 indicates a monopolistic market and 1 indicates a duopolistic market. Let the vector of explanatory variables $\boldsymbol{X'\beta}$ be reduced to just market size. To maintain previous notations, we denote the market size variable as $\boldsymbol{Y}$ and its corresponding parameter as vector of $\boldsymbol{\lambda}$. The formalized notation is then as follows:

$$y = \beta_0 + \delta_N D_N + \boldsymbol{\lambda Y} + \varepsilon \tag{3.26}$$

In this simple example, observed variables are market size $\boldsymbol{Y}$ and market structure D_N, which affect market outputs $\boldsymbol{y}$. The parameter δ_N reflects the difference in market outcomes between monopolistic and duopolistic market structures, given the identical market (as we reduce the demand function $\boldsymbol{X'\beta}$ to $\boldsymbol{\lambda Y}$) and some market shock or unobserved variables captured by $\boldsymbol{\varepsilon}$. Thus, the coefficient δ_N estimates the difference between outcomes "caused solely" by market penetration. In other words, for the same level of other factors, the difference between those market structures is captured by δ_N. Under the zero conditional mean assumption, we yield: (Wooldridge, 2012)

$$\delta_N = E(Q|D_N, \boldsymbol{X'\beta}) - E(Q|D_{N+1}, \boldsymbol{X'\beta}) \tag{3.27}$$

Now it is clear that expression (3.27) is identical to expression (1.8) from the theoretical part. The key here is that the level of demand $\boldsymbol{X'\beta}$ holds the same for all D_N, and the difference δ_N is due to market structure only. To develop a scale-free measurement mirroring ETR, we just divide the corresponding δ_N in a manner such that $ETR = \delta_{N+1}/\delta_N$, again resembling expression (1.9). This very simple idea enables us to evaluate B&R estimation independently and from a distinct methodological viewpoint. This is possible at the price of demanding data collection regarding market outcomes.

To formalize our expectations, we can assume that in the case of perfect competition, the change between market outcomes for individual market structures would be the same for all observed structures. Formally, $\delta_N = \delta_{N+1} = 1$, as individual firms do not exhibit any ability to increase market saturation given the zero-economic profits condition.

Contrary to this, for concentrated markets, we would observe that each entry increases competition, resulting in increased demand saturation. This situation would mean that the level of demand saturation increases with each entry. We can formalize this pattern as $\delta_N > \delta_{N+1} > 1$. Again, our expectations are that with an increase in the number of firms, this relative change would approximate competitive levels defined as a value of 1.0.

3.2.6 Selection bias

In the B&R model, we utilized an ordered probit model for the estimation of the limited dependent variable – the number of firms. This is mirrored in the OLS model, where market structure is represented as dummy variable D_N, where we incorporate vector $\boldsymbol{X'\beta}$. This link is perfectly fine as long as the vector of variables $\boldsymbol{X'\beta}$ sufficiently describes variations in $\boldsymbol{y}(\boldsymbol{Q})$. In other words, as long as the relationship between $\boldsymbol{y}$ and $\boldsymbol{X'\beta}$ is appropriately described, only then are we able to draw inferences about δ_N ceteris paribus. Despite potential inaccuracies stemming from omitted variables, we already know that this relationship suffers from selection bias. (Wooldridge, 2012), (Hicks, 2021), (Abraham et. al., 2007)

To illustrate why our OLS model falls short in this regard, we must delve into the foundational assumptions of the Gauss-Markov assumptions for linear regressions. We summarize them as:

- Assumption 1: The model is correctly specified. That is $y = \boldsymbol{X'\beta} + \varepsilon$.
- Assumption 2: Random sampling. That is, we have random sample size $n, \{(x_i, y_i): i = 1,2, \dots n\}$.
- Assumption 3: Sample variation in explanatory variables.

- Assumption 4: Zero conditional mean, $E(u|x) = 0$.
- Assumption 5: Homoscedasticity, $Var(u|x) = \sigma^2$.
- And others

Our OLS model may fail due to the violation of any of the listed assumptions; however, our particular concern arises from the violation of assumption 2. In general, OLS models violating this assumption are either censored or truncated.

Censored regression occurs when there is a pattern of missingness in variable y_i for certain observations i. In such situations, we can identify the pattern of censored y. On the other hand, truncated regression occurs when a subset of the population is excluded based entirely on y. The pattern of missingness is then typically determined by a threshold of y, above or below which missingness emerges. We omit a detailed description of censored regressions. The idea here is that we usually observe y up to a certain threshold c. In the case of truncated regressions, we observe the explanatory variables for each unit; however, we only observe a subset of y determined by some threshold. Examples of such biases are various surveys that have thresholds, such as identifying gifted and talented students (with the threshold being GTA) or taxpayers (where the threshold is income). Truncated regression models are OLS models, such as (3.25), with respect to the observability of market outcomes if $y_i \leq c_i$, where c_i is the truncation threshold that depends on $\boldsymbol{x}_i$. This implies that we observe $\boldsymbol{x}_i$ only if $y_i \geq c_i$.

One special type of sample selection is incidental truncation. In this case, we always observe all explanatory variables $\boldsymbol{x}$; however, we only observe y_i if another variable z surpasses some threshold c_i, as $y = y_i \ if \ c_i \leq z$. This type of truncation is incidental because it solely depends on another variable. A clear example of incidental truncation is our OLS model. We observe market outputs y_m in market m if there is at least one pharmacy N. Thus, we observe market outcomes if another variable (number of pharmacies) is at least one, as $y = y_i \ if \ N_i \geq 1$.

To address this issue, we employ the Heckman sample selection model, which serves as the standard way to deal with incidental truncation.

3.2.7 The Heckman Sample Selection Model

We aim to estimate model (3.25), which suffers from sample selection bias. For simplicity, we ignore dummy variables. (Hicks, 2021), (Wooldridge, 2012)

$$y^* = \boldsymbol{X'\beta} + \varepsilon, E(u|\boldsymbol{x}) = 0 \tag{3.28}$$

This selection arises from the unobservability of y given markets without any firm, where y^* denotes this selection. In the first step, we estimate this quantity equation to observe its biased estimates. In the second step, we aim to estimate the missingness pattern by employing a probit model to model this latent process. Formally:

$$z^* = \boldsymbol{w}\gamma + u \tag{3.29}$$

Given the requirement that $\boldsymbol{x}$ and $\boldsymbol{z}$ are always observed, we can write $\boldsymbol{X'\beta} = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k$ and $\boldsymbol{Z\gamma} = \gamma_0 + \gamma_1 z_1 + \cdots + \gamma_m z_m$. This standard probit model estimates whether we observe market outcomes y or not. Note, that the probability of observing outcome is identical to the probability of firm entering the market. We observe that:

$$z_i \begin{cases} = 1 & if \quad z_i^* > 0 \\ = 0 & if \quad z_i^* \leq 0 \end{cases} \tag{3.30}$$

Thus $z_i = 1$ means, that we should observe at least one firm on the market given the probability of observing it greater than zero ($z_i^* > 0$). Alternatively, $z_i = 1$ means, that y_i is not censored if:

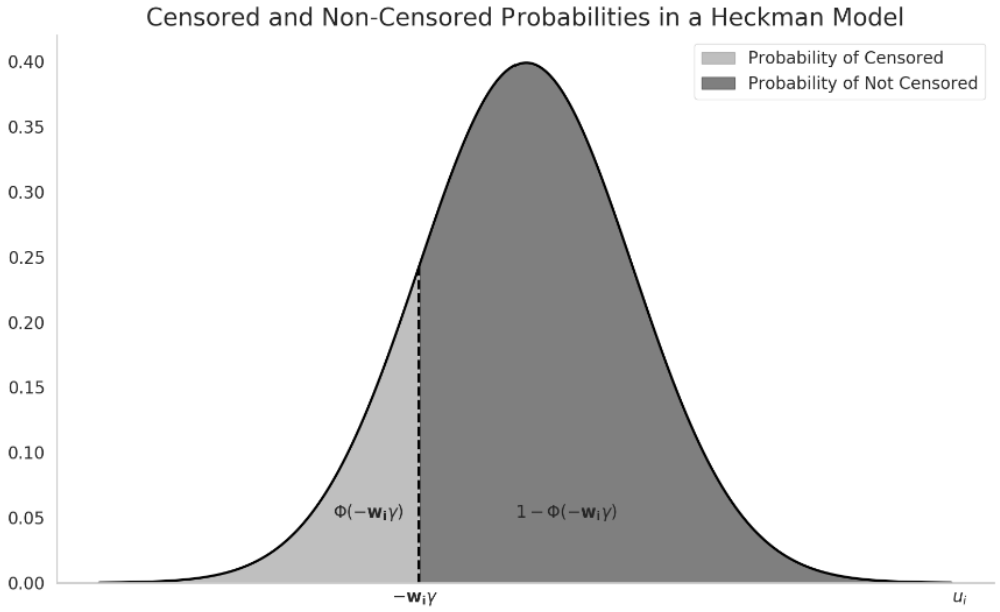
$$u_i \geq -\boldsymbol{w}_i\gamma \tag{3.31}$$

The assumptions here are that the error terms $\boldsymbol{\varepsilon}$ are independent ($E(\boldsymbol{\varepsilon}|\boldsymbol{x},\boldsymbol{z}) = 0$, implying exogeneity of $\boldsymbol{z}$ in (3.28)). Additionally, $\boldsymbol{u}$ is normally distributed with variance 1 and mean 0, ($\boldsymbol{u} \sim N(0, \mathbf{I})$, implying $Var(u_i) = 1$. And that $\phi(\boldsymbol{w}_i\gamma)$ is symmetrical. Then the probability of y_i not being censored is the probability of (3.31) and symmetricity of $\boldsymbol{w}_i\gamma$ implies $1 - \Phi(-\boldsymbol{w}_i\gamma) = \Phi(\boldsymbol{w}_i\gamma)$. Under those assumptions we can write:

$$\Pr(u_i \geq -\boldsymbol{w}_i\gamma) = 1 - \Phi(-\boldsymbol{w}_i\gamma) = \Phi(\boldsymbol{w}_i\gamma) \qquad (3.32)$$

Expression (3.32) implies that the probability of y_i not being censored is the set of errors greater than $-\boldsymbol{w}_i\gamma$. For visualization purposes, we can show (3.32) in Figure 5.

Figure 5: Probabilities in the Selection Mechanism in a Heckman Model



Source: *(Hicks, 2021)*

So far, we can summarize Heckman correction along with some of its assumptions as follows:

$$y_i = y_i^* = \boldsymbol{x}_i\beta + \varepsilon_i\ is\ observed, if\ z_i = 1 \qquad (3.33)$$

Under the assumed joint error distribution of selection (u_i) and equation (ε) model as:

$$\begin{bmatrix} u_i \\ \varepsilon_i \end{bmatrix} \sim Normal\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & \rho \\ \rho & \sigma_\varepsilon^2 \end{bmatrix}\right) \tag{3.34}$$

The summation of first (3.30) and second (3.31) step can be written as:

$$E(y_i|y_i observed) = \begin{cases} E(y_i|z^* > 0) \\ E(y_i|u_i > -\boldsymbol{w}_i\gamma) \\ \boldsymbol{x}_i\beta + E(\varepsilon_i|u_i > -\boldsymbol{w}_i\gamma) \\ \boldsymbol{x}_i\beta + \rho\sigma_\varepsilon \dfrac{\phi\ (\boldsymbol{w}_i\gamma)}{\Phi(\boldsymbol{w}_i\gamma)} \end{cases} \tag{3.35}$$

Form expression (3.35) we see that the condition $E(y_i|y_i observed)$ differs from unconditional mean $\boldsymbol{x}_i\beta$ only if correlation $\rho \neq 0$, since the quantity equation error variance (σ_ε) and Inverse Mill's Ratio $\frac{\phi(\boldsymbol{w}_i\gamma)}{\Phi(\boldsymbol{w}_i\gamma)}$ are strictly positive. Thus, if the $\rho_{\mu,\varepsilon} = 0$, no correlation between error terms is observed, and Heckman correction would be reduced to simple OLS as: $\boldsymbol{X}_i\beta + \rho\sigma_\varepsilon \frac{\phi\ (\boldsymbol{w}_i\gamma)}{\Phi(\boldsymbol{w}_i\gamma)} = \ \boldsymbol{X}_i\boldsymbol{\beta}, if\ \rho = 0$.

The incidental truncation in our model is resolved by estimations correction as $\boldsymbol{X}_i\beta + \rho\sigma_\varepsilon \frac{\phi\ (\boldsymbol{w}_i\gamma)}{\Phi(\boldsymbol{w}_i\gamma)}$. Additionally, we employ MLE method for Heckman correction in our OLS models.

4 Results

This section is dedicated to describing the findings obtained from our analysis of Slovak retail pharmacy market outcomes in 2016 concerning the number of entrants. We present estimates derived from both the B&R and OLS models, aiming to shed light on how market entry influences levels of competition. Through a comparison of these model estimates, we aim to ensure the robustness between both models' estimates.

4.1 Descriptive statistics

In this section, our goal is to utilize descriptive statistics to draw inferences about our datasets. We divide this section into two parts. The first part is dedicated to market data, focusing on the distribution of market structures. Here, we aim to provide insights into the composition of markets in terms of the number of firms operating within them.

The second part is dedicated to explanatory variables for demand, market size and quantity data. We describe characteristics of these variables to understand their distribution and variability across our datasets.

4.1.1 Descriptive statistics for market data

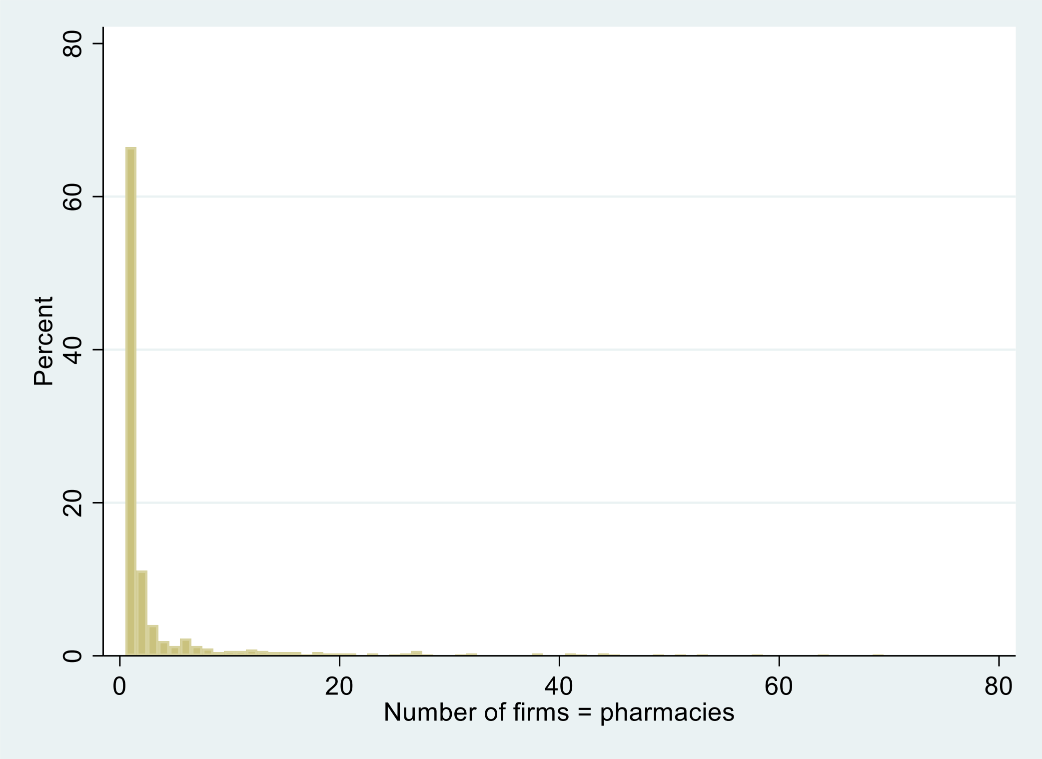
In both the B&R and OLS models, the unit of interest is the municipal market of the Slovak Republic in 2016, defined administratively as a Slovak town. Thus, our models consist of cross-section data. Each market is characterized by continuous demographic and market variables, along with an ordinal number representing the count of pharmacies, ranging from 0 to 69 observed pharmacies in each market. Following the assumptions of the B&R model, we consider these markets to be isolated, with each possessing unique characteristics that do not interact with neighbouring markets. The data table (Table 1.) illustrates the distribution of pharmacy markets in Slovakia based on the number of pharmacies within each market. A significant portion of the markets, comprising 78.82 % of the total, have no pharmacies present. Meanwhile, 14.08 % of markets contain only one pharmacy, with 2.36 % having two pharmacies and 0.85 % having three pharmacies. Additionally, a smaller fraction of markets, representing 0.41 %, have four pharmacies. Markets with five or more pharmacies constitute 3.48 % of the total. Overall, the dataset includes information from 2,927 municipal markets in Slovakia. This pattern of an exponential distribution is expected, as it reflects the natural structure of markets observed in other papers on industrial economics and is consistent with market reality. Thus, the frequency of market structures decreases, following expected distribution. As a result, there are only a few observations in structures with more than 5 firms in each market. If we were to examine these markets using the B&R model (as well as the OLS model), we would have to rely on a small set of observations for more complex market structures, resulting in unrobust estimations. To address this issue and ensure robustness in our models, we decided to set the threshold for analysis to $N = 5$. This threshold allows us to focus our analysis on market structures that are more commonly observed and for which we have sufficient data to draw meaningful conclusions.

Table 1: Frequency of Pharmacy Retail Market Structures in Slovakia (2016)

Structure of pharmacy markets in Slovakia			
Pharmacies in market	Freq.	Percent	Cum.
0	2,307	78.820	78.820
1	412	14.080	92.890
2	69	2.360	95.250
3	25	0.850	96.110
4	12	0.410	96.520
+5	102	3.480	100.000
Total	2,927	100.000	

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic*

Figure 6: Distribution of Pharmacy Markets in Slovakia (Excluding Markets with Zero Pharmacies)



Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

4.1.2 Descriptive statistics for explanatory and quantity data

In the OLS model, the unit of interest is the market outcome realized on the respective market. This measurement is the quantity for over-the-counter medicine, prescribed medicine and sum of the quantity of over the counter (OTC) medicine and the quantity of prescribed medicine (PM). These data were obtained from the Ministry of Health of the Slovak Republic and capture the situation in each market for the year 2016. As the section dedicated to Heckman Selection suggests, we encounter a situation where we do not observe market outcomes, represented by realized quantity, in markets without any pharmacies. This results in fewer observations (observed markets) compared to the situation in the B&R model. Using Heckman selection, as described by Abraham et al. (2007), is a suitable solution for these problems. The Heckman procedure increases the value of the constant in the OLS model and decreases overestimated value of market dummies, which makes them comparable with the estimations of ETR.

In Table 2, we present summary statistics regarding explanatory and quantity variables. All data were obtained from the Statistical Office of the Slovak Republic, except for the quantities that we gathered from Ministry of Health of the Slovak Republic. All data are from the year 2016. Thus, out dataset consists of cross-section data. The variable for quantity was calculated as the sum of over-the-counter (OTC) and prescribed medications realized in each pharmacy of a market.

The market size variable is represented by the household variable, which indicates the number of households within the market. Variables productive, old, and young represent the number of people in each respective category in 2016. The variable unemployed signifies the percentage of unemployed people in the market and the variable wages refer to the average wage in each market in 2016. Those variables collectively represent demand shifters.

Extreme values in the dataset are caused by villages such as Šarbov, Príkra, Olšinov, Veľké Borové, Ondavka, and others, where some of the observations have zero value. For the purpose of data normalization and resolving these extreme values, we use a logarithmic model. This approach will result in losing some observations with extreme values but will provide us with easy-to-explain results, as the independent variables represent the elasticity of the number of firms. Another important distinction is the parameter of the number of households. We used this variable as the measure of market size instead of the number of population (as Abraham et. al. (2007) or Bresnahan and Reiss (1991)) because the statistical significance for households was much higher than that for population in all conducted models.

Table 2: Summary Statistics of Pharmaceutical Market Variables in Slovakia (2016)

Descriptive Statistics

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Q	All quantity realized	2924	47919.27	263328.34	0	4368241
quantity OTC	Realized OTC quantities	2924	20063.233	123521.88	0	2167858.3
quantity medication	Realized PM quantities	2924	27856.037	142198.93	0	2271229.5
households	Number of households	2924	362.001	1326.603	1	25341
wage 2016	Average market wages in 2016	2924	858.537	113.395	658	1486
old 2016	Number of old people	2924	278.446	928.648	1	14414
young 2016	Number of young people	2924	287.228	826.417	0	14795
productiv 2016	Number of people in productive age	2924	1292.255	4060.721	5	74934
unem 2016	Percentage of unemployed people	2924	11.221	5.702	3.29	24.58
wage pharmacists	Pharmacists' wages (fixed costs)	2924	13941.455	1736.915	11166.87	17589.768

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

4.2 Results of B&R model

In Table 3, we present the results from the ordered probit model, whose estimates are used in the calculations of the B&R entry model. The signs of the estimates are reasonable and most of them consistent with theory and expectations. Note that the direct comparison of estimates is heuristic, as we should compare marginal values.

The parameter for market size (number of households) is strongly positive and statistically significant. This indicates that the expected probability of the number of pharmacies increases with a rising number of households. The demographic variable for the number of old people is positive and statistically significant, suggesting that the presence of more elderly individuals in the market leads to a higher probability of the expected number of firms in the market. Interestingly, the parameter estimate for unemployed people in the model is unexpectedly positive and statistically significant. This finding suggests that there is a notable association between the percentage of unemployed market inhabitants and pharmaceutical market structures in Slovakia. One potential explanation for this result could be attributed to the relatively poorer health conditions experienced by unemployed individuals compared to those who are employed or have a stable income.

The negative sign of the estimate in the category of young population is consistent with the expectation that in this demographic category, the demand for pharmacy products is not as large as in other demographic categories. The negative estimator for pharmacists' wages is expected, given that this variable is treated as a fixed cost parameter in the model. As fixed costs increase, such as pharmacists' wages, the profitability of operating pharmacies may decrease, leading to a negative association between pharmacists' wages and number of entrants. This result is in line with economic theory, which suggests that higher fixed costs can constrain profitability. The low statistical significance of parameters for wages and the demographic parameter of the population in the productive age group can be attributed to their variance in relation to market structures. Regarding wages, while higher wages typically have a positive effect on pharmacy profits and individual purchasing power, they may also indirectly affect health and, consequently, promptness to purchase pharmaceutical products. This dual effect complicates the relationship between wages and market structures, resulting in lower statistical significance. Additionally, it's important to note that the model suffers from imperfect multicollinearity, despite having a sufficiently high number of observations. This multicollinearity can lead to reduced precision of the model's predictions. Therefore, the results of the model should be interpreted with this limitation in mind, acknowledging the potential for inflated standard errors and less reliable estimations.

The most important estimations in this model are the cut values. These estimations are statistically significant, and the coefficient of those estimates rises with respect to the number of firms in the market. As expected, their change is smaller with each entry. Interpreting the estimations of the cut values can be challenging. Essentially, the value of cut N represents the probability density of market characteristics that we observe at least N firms. The expected pattern of cut values mirrors the expectations of entry thresholds. These values rise, representing the expected increase in market size necessary to support new entrants.

However, the change of each cut (the market size) is smaller than the previous one as more firms enter.

This expected pattern mirrors the assumption about convergence to the market with perfect competition. In other words, as more entrants enter the market, the relative change in market outcomes becomes less prevalent, eventually reaching perfect competition, where the change is either negligible or equal to 1.0.

Our estimates align with this expected pattern. As more pharmacies enter the market, the cut values rise, but the rate of change becomes smaller with each entry. For example, the first cut value (cut1) with a value of 9,1 represents the point in the probability density of the ordered probit model where there exists the probability of observing at least the monopoly structure in the markets. Similarly, the second cut value (cut2), with a value of 10.9, represents the threshold at which the level of market size can support two pharmacies, achieving the necessary probability to be observed. The gap between these two points is approximately 1.8. This gap is ever smaller with each entry. In the case of the final fifth cut (cut5), the gap is approximately 0.3. The importance of these estimations lies in the fact that all of them enter the equation for entry threshold.

Table 3: Ordered probit regression model of market structures

N_pharmacy	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]
ln_households	1.319	.086	15.40	0	1.151	1.487
ln_wage_2016	.426	.362	1.18	.239	-.283	1.134
ln_old_2016	.371	.176	2.11	.035	.026	.716
ln_young_2016	-.232	.17	-1.36	.174	-.565	.102
ln_productiv_2016	.374	.314	1.19	.234	-.242	.989
ln_unem_2016	.311	.087	3.57	0	.14	.481
ln_wage_pharmacists	-.555	.268	-2.07	.039	-1.082	-.029
cut1	9.096	3.249			2.729	15.464
cut2	10.868	3.252			4.494	17.242
cut3	11.704	3.256			5.322	18.086
cut4	12.206	3.26			5.816	18.595
cut5	12.504	3.262			6.11	18.899

Mean dependent var	0.405	SD dependent var	1.043
Pseudo r-squared	0.532	Number of obs	2918
Chi-square	2278.082	Prob > chi2	0.000

*** p<.01, ** p<.05, * p<.1

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

4.3 Heckman selection model

Tables 4.-6. present the results from both the OLS and Heckman selection OLS models. In these models, we regress the natural logarithm of the quantity (OTC, PM, and their respective summary) realized in the market on the natural logarithm of all parameters from the B&R model, except for pharmacist's wage, which serves as the exclusion restriction in the Heckman selection model.

The rationale behind including pharmacist's wage is that it serves as the best available proxy for fixed costs. The inclusion of pharmacist's wage is motivated by the assumption that fixed costs primarily influence firms' decisions to enter the market rather than impacting the realized quantities once firms have entered.

Within these models, the market structure dummy variable denotes the structure of markets based on the number of firms present in the market, ranging from one to five or more firms, with the fifth firm serving as a dummy for five or more firms in the market.

In the Heckman Selection models, estimates for markets with monopolies are not available. However, this limitation does not impede our analysis, as our primary focus lies in understanding the relative changes in market outcomes with the entry of new firms. Starting with the simple OLS model results (first column), it's noted that these outcomes are subject to selection bias due to the unobservability of quantity data for markets without any pharmacies. While the estimate for market size is statistically significant, estimates for demand shifters are largely insignificant. However, these demographic variables are treated as control variables, given our main interest in estimating dummy variables representing market structure. Notably, estimates for the number of pharmacies (denoted as δ_N) are significant and increase with each additional firm with diminishing magnitudes. Furthermore, the small parameter estimate for the constant term in the model is attributed to overestimation caused by selection bias. Consequently, estimates for δ_N are also overestimated and not directly comparable with Entry Threshold Ratios (ETRs). To address this challenge, the Heckman correction is applied to adjust these estimates.

Corrected OLS estimates (second column) are regressed using the same parameters as the simple OLS model. This leads to an increase in the value of the constant parameter and a subsequent decrease in the estimation of dummy variables representing market structure, aligning with our research objectives. In terms of the model estimates, most variables exhibit insignificance except for market size and the dummy variables representing market structure. All estimated dummy variables are statistically significant. However, there is a slight decrease in the significance of the estimator for markets with four pharmacies.

Table 4: OLS and Heckman selection results for all outcomes of pharmacy markets

	Linear regression		**Heckman selection**	
ln_Q	Coef.	St.Err.	Coef.	St.Err.
ln_households	.09	.025	.846	.125
ln_wage_2016	-.402	.117	-1.344	.442
ln_old_2016	.057	.043	.029	.246
ln_young_2016	.009	.037	.422	.254
ln_productiv_2016	-.044	.069	-.408	.446
ln_unem_2016	-.008	.029	-.05	.12
number of firms				
δ_1	9.973	.036	.	.
δ_2	11.042	.077	0.578	.16
δ_3	11.699	.123	0.845	.257
δ_4	12.312	.175	1.275	.367
δ_{5+}	13.185	.081	1.288	.254
Constant	2.31	.842	14.84	3.405

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

In this model, we regressed the natural logarithm of the quantity summary on the natural logarithm of the market size parameter, the number of households, and the natural logarithm of demand shifters (average wages, number of people in old, young, and productive age and percentage of unemployed individuals). Additionally, parameters $\boldsymbol{\delta_N}$ represent dummy variables for respective market structures. In the case of linear regression, control estimates are largely insignificant with respect to parameters for households and wages. In the case of Heckman correction, we observe a similar situation with control variables. As already described, the point of interest is the estimation for dummy variables. In both models, dummy estimations are strongly significant and increase with reducing magnitude. This suggests an approach towards competitive conduct. By comparing estimated standard errors, we observe a drop in statistical significance in the dummy for markets with four firms.

Table 5: OLS and Heckman selection results for OTC outcomes of pharmacy markets

	Linear regression		**Heckman selection**	
ln_OTC	Coef.	St.Err.	Coef.	St.Err.
ln_households	.06	.039	.857	.199
ln_wage_2016	-.304	.179	-.984	.695
ln_old_2016	.049	.065	-.216	.387
ln_young_2016	.025	.057	.627	.399
ln_productiv_2016	-.001	.105	-.166	.7
ln_unem_2016	.011	.044	-.006	.189
number of firms				
δ_1	8.385	.056	.	.
δ_2	9.945	.117	.984	.252
δ_3	10.686	.187	1.248	.405
δ_4	11.219	.266	1.497	.577
δ_{5+}	12.196	.123	1.472	.402
Constant	1.449	1.281	9.127	5.353

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

In table 5 we summarized OLS and Heckman model for OTC medicine in framework identical to the previous one. Here, among control variables only market size parameter remains significant. For dummy variable estimates, we observe significant approximation of competitive conduct. From initially high difference in demand saturation between monopolistic and duopolistic markets to relatively insignificant difference between markets with four firms and fully penetrated markets. Again, dummy estimates remain significant in this model as well. However, statistical significance of constant lack levels of dummies estimations.

Table 6: OLS and Heckman selection results for prescribed medicine outcomes of pharmacy markets

	Linear regression		**Heckman selection**	
ln_medication	Coef.	St.Err.	Coef.	St.Err.
ln_households	.095	.027	.823	.133
ln_wage_2016	-.458	.123	-1.444	.467
ln_old_2016	.047	.045	-.044	.26
ln_young_2016	-.013	.039	.181	.268
ln_productiv_2016	-.016	.072	-.107	.471
ln_unem_2016	-.004	.03	-.004	.127
number of firms				
δ_1	9.567	.038	.	.
δ_2	10.6	.08	.542	.17
δ_3	11.167	.129	.734	.272
δ_4	11.854	.183	1.213	.388
δ_{5+}	12.654	.085	1.209	.269
Constant	2.635	.88	14.779	3.6

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

Model captured in table 6 represents OLS and Heckman results for prescribed medicine. In this model control estimate for market size and wages is statistically significant. The same goes for all dummies and constant. Results for dummy estimates are almost the same as for model with all outcomes. Thus, we observe pattern of approximation towards competitive conduct with slight deviance, both in estimates and significance, in markets with four firms.

To fully present respective models, we attach tables with results from linear regression and Heckman selection model in Attachments, where λ, ρ and σ_ε from Heckman models can be found.

4.4 Relative Threshold Ratios

Table 7 presents the ETRs derived from both the B&R model and respective selection OLS models. In the B&R model, these estimations are computed in the same manner as ETRs, but the threshold for monopoly (s1) is consistently used as the denominator in the ETR equation. In the selection models, these estimates are computed from the dummy variable for market structure obtained from the Heckman model, raised to the power of Euler's number, given that the selection model is a log-log model.

The results demonstrate a considerable similarity between the two sets of estimates, except for the expected thresholds for markets with four firms. In this case, there are significant differences between the results obtained from the B&R model and those derived from selection OLS models, deviating OLS estimates from both expected patter and B&R trajectory, with except for OTC model. In case of OTC model, we observe slightly inflated estimates consistent with expectations, as estimations for all observed market structures approximate competitive levels. These results are further illustrated in Table 6, which presents the casual Threshold Ratios from both models.

Table 7: Relative Threshold Ratios from B&R and Selection model

	Relative Threshold Ratios			
Ratio	B&R	OLS selection (Q)	OLS selection (OTC)	OLS selection (PM)
s2/s1	1.916	1.782	2,675	1,719
s3/s1	2.407	2.328	3,483	2,083
s4/s1	2.641	3.579	4,468	3,364
s5+/s1	3.313	3.625	4,358	3,350

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

Table 6 displays the Entry Threshold Ratios (ETRs) derived from both the B&R model and respective selection OLS models. In the B&R model, these estimates were computed as the ratio between estimated minimal market size with N firms and the minimal market size with monopolist market structure. In case of OLS models as ratio between market structure dummies for respective market quantities.

As previously explained, this ratio indicates how market outcomes change concerning the number of firms in the market, illustrating the impact of market structure on competition and its convergence towards the state of perfect competition. In the OLS models, the interpretation of results is similar, focusing on how market outcomes (specifically, realized prescribed medication, over the counter medications and their respective summaries) change with each entry into the market. The key difference lies in the fact that the B&R model indirectly estimate firm profits, whereas the OLS model primarily focuses on demand saturation. Consequently, both models calculate these estimates using different methods and approaches, with the number of firms being the only key common parameter. The vector of market size and demographic parameters serve as control variables for estimated slope in the OLS models. Despite these differences, the estimated results are highly similar. By applying assumptions about the fixed nature of average and fixed costs, the interpretations of the ETRs derived from both models remain consistent.

The expected trajectory of Entry Threshold Ratios (ETRs) concerning the increasing complexity of market structures tends towards a state of perfect competition characterized by a ratio value of 1.0. This trajectory is observable in both the B&R model and the OLS selection models for all regressed quantities. In the B&R model, the transition from a monopolistic to a duopolistic market is reflected in an ETR value of 1.91, which decreases to 1.26 by third entrant. And finally, to 1.25 in the case of fully penetrated markets with five and more firms. This trajectory is disrupted in estimations 1.1 for markets with four firms. Similarly, in the OLS selection models, this trajectory is even more pronounced.

For all quantities model the ETR value decreases from 1.78 for markets with duopolistic structure, to 1.3 for markets with three firms and to 1.01 in the case of five and more firms. These results strongly suggest that the pharmacy market in Slovak Republic is not concentrated, and as the number of firms increases, it tends to converge towards a state of perfect competition. For OTC model, this change is even more prominent as the estimated change in demand saturation changes from 2.68 to 1.3 by entry of third firm. In markets with fourth entrant, estimated ETR is 1.28 and changes to ~1, as more than five firms enter the market. Suggesting strong advance to competitive conduct. The same holds true for model regarding prescribed medicine, where duopolistic ETR change from 1.7 to 1.2. Subsequent entries lied to competitive level of ~1, again suggesting deepening competition intensity.

Table 8: Threshold Ratios from B&R and Selection models

		Relative Threshold Ratios			
	Perfect competition	B&R	OLS selection (Q)	OLS selection (OTC)	OLS selection (PM)
s2/s1	1,00	1,916	1,782	2,675	1,719
s3/s2	1,00	1,257	1,306	1,302	1,212
s4/s3	1,00	1,097	1,537	1,283	1,614
s5+/s4	1,00	1,254	1,013	0,975	0,996

Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

The main deviation from the estimations of both B&R and OLS models and the violation of the expected trajectory are the results for markets with four pharmacies. Except for OTC model, which exhibit expected trajectory of all dummy estimations Interestingly, these results diverge in opposite directions between the B&R model and the OLS model. There are several potential reasons for these unexpected estimations.

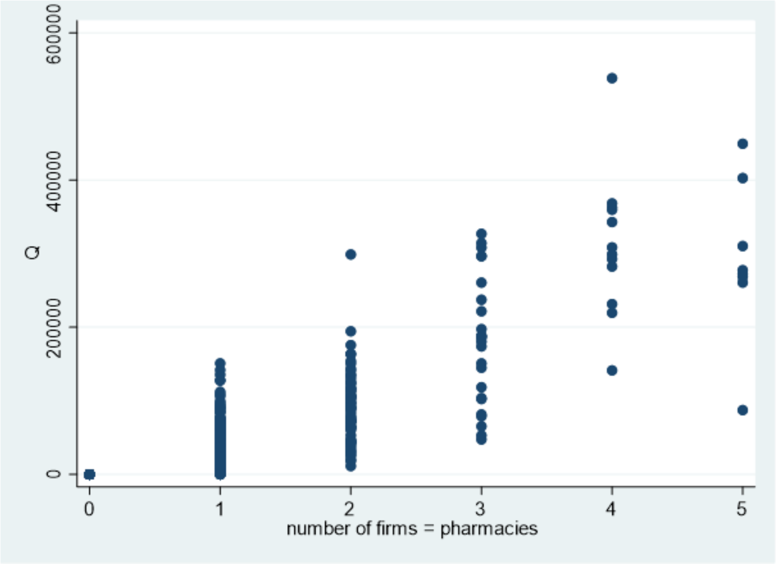
From a theoretical perspective, one factor could be the definition of market boundaries based on administrative borders. For example, markets like Bojnice, which are near larger cities like Prievidza, or Svit near Poprad, may experience influences from neighboring markets that are not captured by the administrative definition. Similarly, districts like Sídlisko Ťahanovce in Košice city and Šahy, bordering Hungary, may also be subject to external influences that affect market dynamics.

Second factor streaming from OTC estimations is the exogeneity of prescribed medicine quantities on the retail pharmacies behaviour. The question here is whether a pharmacy itself can influence realized quantities of prescribed medicine. As answering this question is the subject of another urgent issue, we do not investigate it further in this thesis and we simply assume that this exogeneity does not hold significance here.

Another, factor to consider is the possibility of cost changes for firms operating in these markets. Proximity to larger markets could lead to significant fluctuations in costs, impacting market behaviour and outcomes. While theoretical explanations offer insights, further exploration of the data may reveal additional factors contributing to this deviation. By delving deeper into the data, we may uncover more precise explanations for the unexpected results observed in markets with four firms.

Figure 9 depicts the quantity realized on the market with respect to the number of pharmacies in each market. Interestingly, markets with four firms exhibit quantities realized that are, on average, almost as high as those in markets with five firms. Notably, outliers such as Kráľovský Chlmec (bordering a city) and Vajnory in Bratislava (part of Bratislava city) significantly impact the data. While manually dropping these outliers may improve estimates, this approach is not advisable as it does not address the underlying issue. The primary reason for this situation is the limited number of observations for both markets. This limitation is not applicable to markets with fewer than four firms or more than five firms. The difference in the direction of change (positive/negative) between the two models stems from their distinct approaches. B&R calculates the ratio between two markets (S_N/S_{N+1}), with a larger result for the $N+1$ market yielding a smaller estimated result. Conversely, in the OLS approach, where ETR is calculated as δ_N/δ_{N-1}, a larger value of the dependent variable y_N leads to a greater estimate of β_{D_N} and constant ceteris paribus.

Table 9: Relationship between quantity of all pharmaceutical products consumed and number of pharmacies in markets



Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

Back to the results from table 8, the estimates for the ETRs in duopolistic markets differ slightly between the B&R and OLS models, with the B&R estimated value of 1.92 and OLS estimations of 1.78 for all quantities and 1,72 for prescribed medicine. Apart from ETR of 2.7 in the OTC medicine, where we notice a higher change in the level of demand saturation. B&R estimates imply that markets must be at least 1.92 times larger (in terms of representative consumers) than markets with monopolistic structures to support the entry of a second firm. In the OLS model, the interpretation suggests that the entry creating duopolistic market structure increases market outputs by 178% in case for all quantities, 172% for prescribed medicine and 268% for OTC quantities, as the estimated ETR has value of 2,68. This estimation implies very strong shift between market outcomes suggesting extensive change in competition.

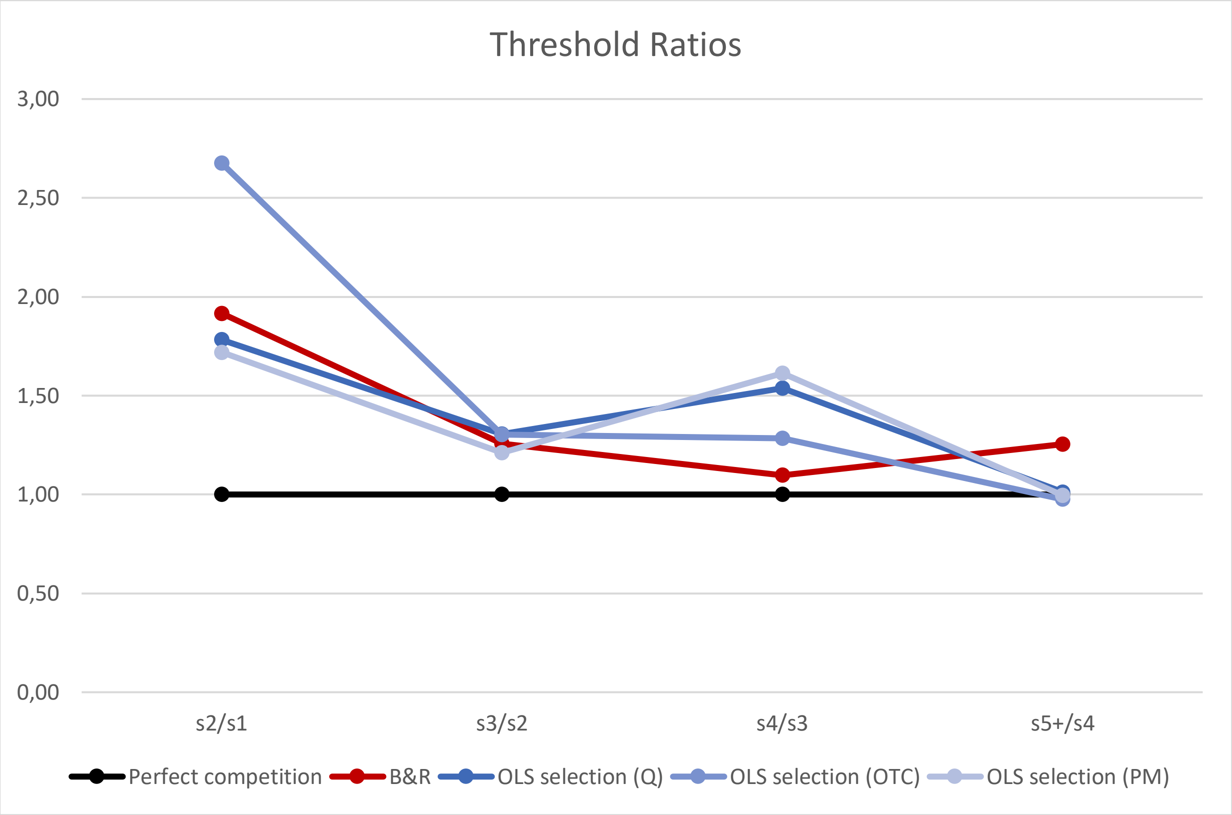
Despite the differences in these estimates, they provide a good approximation to each other. Making the estimates more reliable and robust. However, both estimates are still far from reflecting perfect competition (approaching an estimate close to 1.0). Yet, the primary focus of our model is not the actual state of competition but rather the relative change with each market structure.

For triopoly markets (B&Rs s3), the market needs to be at least 1.26 times larger than one with just two firms to accommodate the entry of the third firm. Additionally, the presence of the third firm increases market output on average by ~131% in case of OTC and all quantities and by 121% in case of prescribed medicine. These estimates indicate a gradual approach toward the state of perfect competition, where minimal discrepancies within estimations are observed. For markets with four firms, the situation becomes somewhat complicated.

As previously explained, the results from the B&R and OLS models differ both from each other and from the expected trajectory. The B&R model estimates are nearly 1.0, suggesting that the market has achieved a state of perfect competition with just four firms. However, the OLS estimate of 1.53 and 1.64 are significantly higher. This indicates that the presence of four pharmacies in the market increases all realized quantities by 153% and prescribed quantities by 162%. This is not true for case of OTC quantities, as their estimation is consistent with both B&R model and expected pattern. Once more, these estimates should be considered with attention to their statistical significance in mind.

Finally, the estimations for fully penetrated markets (markets with five or more pharmacies) are 1.25 in the B&R model and ~1 in all OLS models. A B&R estimator of 1.25 indicates that for a new firm to enter the market, the market size must be at least 1.25 times larger than a market with four firms. On the other hand, OLS results suggest perfect competition on this market. All results are presented in Figure 7, where the trajectories of ETRs from the B&R model (red line) towards the state of perfect competition (ETR = 1.0) can be observed, with ever-smaller relative changes. In the case of the OLS models (blue lines), these estimates are even more consistent with expectations. Especially for OTC quantities model, where from initially large ETR of 2.7 we observe constant approximation to the state perfect competition, even for markets with four firms, ending at competitive conduct value of ~1.

Figure 7: Relative Changes in Market Competition: Comparison of ETRs from B&R and OLS Models



Source: *Own calculation based on the data from Statistical Office of the Slovak Republic and the Ministry of Health of the Slovak Republic*

5 Discussion

In this thesis we employed entry model based on the Bresnahan and Reiss (1991) approach and OLS model of quantities to enhance the precision of our inferences about *how market outcomes in the pharmacy retail market changes with respect to the market structure*.

Analysis of both models suggests that once entry occurs, the pharmacy retail markets in Slovakia are transitioning towards a state of perfect competition. Furthermore, the consistency between both estimates highlights the robustness of the B&R model. This assertion is supported by the OLS model estimates, which serve as a robustness check for the widely employed B&R methodology.

Our findings regarding the transition towards competitive conduct in Slovak retail pharmacy markets align with those of other papers investigating the same issue, namely Labáj et al. (2018) and Kališ (2023). Moreover, our novel methodological approach to robustness verification may enrich the literature on empirical entry models.

There are several downsides in our analysis. The first one is the definition of markets. Our markets are defined administratively, by the town's borders. Unlike (Labáj, et. al., 2017) we did not employ any special model to enhance market definition, thus stretching our assumptions about market isolation. Even in original paper by Bresnahan and Reiss (1991) the markets are assumed to be isolated as they investigated geographically remote markets in USA.

Second drawback regarding methodology is the violation of Gauss-Markow theorem's assumption regarding zero conditional mean in OLS models. As Abraham et. al. (2007) proves, this assumption is violated in our case. Even though their paper is much more complicated than our one, we can show how endogeneity affects our analysis. We specified our OLS model in expression (3.25). By expanding it as:

$$y = \boldsymbol{X}'\boldsymbol{\beta} + \sum_{i=1}^{N} \delta_N D_N + \varepsilon$$

Here, we can rewrite this expression as $Q_N = Y\lambda + \boldsymbol{X}\beta_X + \delta_N + \varepsilon_Q$, where $\varepsilon_Q = \varepsilon_S + \varepsilon_d + \varepsilon$. Here, dependent variable y is a quantity Q for market structure consisting of N firms

Q_N, s and d donates market size and demand of representative consumer, λ is estimated parameter of market size Y. β is parameter of demand shifters $\boldsymbol{X}$, δ_N is dummy for market structure N. We see, that in markets where the realised quantities $\boldsymbol{Q_N}$ are higher, also unobservable $\boldsymbol{\varepsilon_Q}$ are higher. This would not be problem, unless we expect that in markets with higher $\boldsymbol{Q_N}$ the number of firms tend to be also higher. Thus $E\left(\varepsilon_Q, \varepsilon_S, \varepsilon_d \middle| \delta_N\right) \neq 0$, which violates assumption about $E(\varepsilon|\boldsymbol{x}) = 0$. This leads to biased estimation of δ_N in a way that the effect of competition on quantity looks greater than it in fact is. Under this perspective the discrepancy between B&R and OLSs estimates for fully penetrated markets may be manifestation of this bias caused by endogeneity.

Another crucial theoretical problem within the OLS model methodology that we aim to address is the correlation between the quantity of prescribed medicine and competition in retail pharmacy markets. In this thesis, without delving into a deeper analysis of this pressing issue, we question whether the saturation of demand for prescribed medicine is exogenous or endogenous concerning the behaviour of retail pharmacies. Put simply, can a retail pharmacy influence the quantity of prescribed medicine? Our analysis within the OLS models implicitly assumes complete endogeneity of this relationship. Despite the potential criticism of this assumption, our analysis of over-the-counter quantities, with some implicit assumptions, does not suffer from this shortcoming.

We envision the potential of applying discrete factor approximation developed by Thomas Mroz in our thesis to tackle the challenge of underlying endogeneity in OLS models and to achieve separational identification of the driving forces of competitiveness within market structures. This method has been previously utilized in a paper by Abraham et al. (2007), in which they investigate similar issue on hospital data in the USA. Additionally, this method allows to relax some normal distribution assumptions.

6 Conclusion

In this thesis, we explored the impact of pharmacy retail market structure on market outcomes. We employed two approaches for improve the accuracy of our investigation. Firstly, we based our methodology on the entry models by Bresnahan and Reiss (1991), utilizing easily observable information on market size and demand shifters (e.g., age structure, average wages). Secondly, we utilized the Ordinary Least Squares selection models, where the market-level outcome was measured as quantity of prescribed medicine, over-the-counter medicine, and their sum.

Upon comparing the two models estimates, our analysis indicates that pharmacy retail markets in Slovakia are moving towards a state of perfect competition once entry occurred. This suggests a lack of significant inefficiencies or concentration in the market. Specifically, markets with four firms (in the case of B&R) and markets with five firms (in the case of the OLS models) exhibit Entry Threshold Ratios very close to one. This implies that there are no substantial changes in outcomes concerning shifts in market structure. In simpler terms, the addition of a new entrant does not bring substantial benefits to the efficiency of these markets.

It's worth noting that markets with these structures may be inaccurately estimated due to their limited prevalence. Furthermore, our OLS selection models serve as a robustness analysis for the commonly used B&R model. The B&R model is popular, primarily due to its low data requirements. Thanks to our collaboration with the Ministry of Health of the Slovak Republic, we were able to observe market-level outcomes. The results of this robustness check align closely with the overall findings of the more commonly used B&R model, affirming the reliability of such conclusions. These findings indicate that the Bresnahan and Reiss model can serve as a reliable and robust initial framework for analysing efficiencies in local markets when data availability is limited.

7 Resumé in the Slovak language

Táto diplomová práca sa zaoberá lekárenským trhom na Slovensku v roku 2016. Hlavným cieľom práce je preskúmať vzťah medzi zmenami v trhovej štruktúre a konkurencií na trhu. V práci využívame dva metodologické prístupy, prvý prístup je model vstupov založený na metóde Bresnahana a Reissa (1991). Druhým je metóda najmenších štvorcov s umelými premennými. Odhady oboch prístupov konfrontujeme za účelom overenia robustnosti oboch odhadov. Výsledky naznačujú zvyšovanie intenzity konkurencie s každým ďalším vstupom na trh. Inovatívnosť tejto práce spočíva v použití metódy najmenších štvorcov s umelými premennými za účelom overenia robustnosti odhadov získaných metódou B&R.

Metóda B&R využíva „usporiadaný probitový model“[2] na odhad vstupných hraníc na základe podmienky nulového zisku. Takto odhadnuté vstupné hranice predstavujú minimálnu veľkosť trhu pre vstup istého počtu firiem na trh. Za účelom vytvorenia bezrozmerného ukazovateľa dynamiky zmeny konkurencie (vzhľadom na vstup ďalšej firmy) využívame koncept pomeru vstupných hraníc. Na základe týchto hodnôt potom vyvodzujeme závery o zmenách v dynamike konkurencie vzhľadom na jednotlivé trhové štruktúry.

Druhá metóda použitá v našej práci je metóda najmenších štvorcov s umelými premennými. V tomto metodologickom prístupe je vysvetľovanou premennou počet realizovaných množstiev voľnopredajných liečiv, liečiv na predpis a súčet týchto množstiev. Vysvetľujúcimi premennými sú vektor premenných pre odhad dopytu reprezentatívneho spotrebiteľa a premenná pre veľkosť trhu. Závery o dynamike konkurencie sú založené na pomere odhadov umelých premenných pre jednotlivé trhové štruktúry.

Pri uplatnení tohto prístupu však narážame na problém so selekciou vzorky. Konkrétne na problém s „náhodne orezanou vzorkou“[3], kde je selekcia podmienená na

---

[2] Z anglického ordered probit model.

[3] Z anglického “ incidental truncation selection”

modeli nezávislou premennou, konkrétne prítomnosťou lekárne na trhu. Za účelom odstránenia tohto problému uplatňujeme Hackmanovu korekciu.

V oboch modeloch je trh definovaný ako obec na Slovensku v roku 2016. Každý trh je charakterizovaný prierezovými údajmi o množstvách realizovaných liečiv, trhovej štruktúre a demografickými údajmi. Tieto pozostávajú z počtu domácností, priemerných miezd, počtu obyvateľov v jednotlivých demografických kategóriách (osoby v predproduktívnom, produktívnom a poproduktívnom veku), podielu nezamestnaných a údajov o mzdách farmaceutov.

Výsledky porovnania odhadov oboch modelov naznačujú, že trhy lekární na Slovenku sa s každým ďalším vstupom približujú k úrovni dokonalej konkurencie. To naznačuje neprítomnosť výraznej neefektívnosti a koncentrácie na tomto trhu. Výsledky modelu B&R odhadujú dosiahnutie úrovne dokonalej konkurencie pri trhoch so štyrmi firmami. Pri metóde najmenších štvorcov je táto úroveň odhadnutá pre trhy s piatimi firmami.

Zároveň podobnosť odhadov oboch použitých metód dokazuje ich robustnosť. Toto zistenie naznačuje, že model B&R môže slúžiť ako spoľahlivá metóda pre analýzu efektívnosti lokálnych trhov.

8 Biography

Abraham, J. M., Gaynor, M., & Vogt, W. B. (2007). Entry and competition in local hospital markets*. The Journal of Industrial Economics, 55(2), 265–288. https://doi.org/10.1111/j.1467-6451.2007.00311.x

Balmer, R. E. (2013, December 19). Entry and competition in local newspaper retail markets - when two are enough. Entry and Competition in Local Newspaper Retail Markets - When two are enough. https://mpra.ub.uni-muenchen.de/54079 /1/MPRA_paper_54079.pdf

Berry, S. T. (1992). Estimation of a Model of Entry in the Airline Industry. Econometrica, 60(4), 889–917. https://doi.org/10.2307/2951571

Berry, S. T. (1992). Estimation of a Model of Entry in the Airline Industry. Econometrica, 60(4), 889–917. https://doi.org/10.2307/2951571

Bresnahan, T. F., & Reiss, P. C. (1991). Entry and competition in concentrated markets. Journal of Political Economy, 99(5), 977–1009. https://doi.org/10.1086/261786

Greene, W. H. (2003). Econometric analysis. Prentice-Hall.

Hicks, R. (2021). ECON407 Cross Section Econometrics. The Heckman Sample Selection Model. https://econ.pages.code.wm.edu/407/notes/docs/index.html

Kališ, R. (2023). Economics of the Pharmacy Retail Market. Wolters Kluwer.

Kališ, R. Economics of the Pharmacy Retail Market: Structure, Competition, and Efficiency. 2023, Wolters Kluwer

Lábaj, M., Morvay, K., Silanič, P., Weiss, C., & Yontcheva, B. (2018). Market structure and competition in transition: results from a spatial analysis. Applied Economics, 50(15), 1694-1715.

Lábaj, M., Silanič, P., Weiss, C., & Yontcheva, B. (2018). Market structure and competition in the healthcare industry: Results from a transition economy. The European Journal of Health Economics, 19(8), 1087–1110. https://www.jstor.org/stable/48699959

Lisý, J.et. al. (2016). Ekonómia. Wolters Kluwer.

Mazzeo, M. J. (2002). Product Choice and Oligopoly Market Structure. The RAND Journal of Economics, 33(2), 221–242. https://doi.org/10.2307/3087431

Newhouse, J. P. (1987). Health Economics and Econometrics. The American Economic Review, 77(2), 269–274. http://www.jstor.org/stable/1805462

Nicholson, W., & Snyder, C. (2008). Microeconomic theory basic principles and extensions (10th ed.). Thomson.

Reiss, P. C., & Spiller, P. T. (1989). Competition and Entry in Small Airline Markets. The Journal of Law & Economics, 32(2), S179–S202. http://www.jstor.org /stable/725595

Schaumans, C., & Verboven, F. (2008). Entry and regulation: Evidence from Health Care Professions. The RAND Journal of Economics, 39(4), 949–972. https://doi.org/10.1111/j.1756-2171.2008.00045.x

Schaumans, C., & Verboven, F. (2011). Entry and competition in differentiated products markets. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1963113

Smatana, M., Pažitný, P., Kandilaki, D., Laktišová, M., Sedláková, D., Palušková, M., Ginneken, E. van, & Spranger, A. (2016). Slovakia: Health system review. Health Systems in Transition, 18(6), 1-210.

Wooldridge, J. M. (2012). Introductory Econometrics: A Modern Approach by Wooldridge (5th ed.). South-Western, Cengage Learning.

9 Attachments

Linear regression for all outcomes

ln_Q	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln_households	.09	.025	3.54	0	.04	.14	***
ln_wage_2016	-.402	.117	-3.42	.001	-.632	-.172	***
ln_old_2016	.057	.043	1.33	.184	-.027	.141	
ln_young_2016	.009	.037	0.25	.8	-.063	.082	
ln_productiv_2016	-.044	.069	-0.63	.528	-.18	.092	
ln_unem_2016	-.008	.029	-0.26	.792	-.064	.049	
number of firms							
1	9.973	.036	273.32	0	9.902	10.045	***
2	11.042	.077	143.82	0	10.891	11.192	***
3	11.699	.123	95.02	0	11.457	11.94	***
4	12.312	.175	70.46	0	11.969	12.655	***
5	13.185	.081	162.65	0	13.026	13.344	***
Constant	2.31	.842	2.74	.006	.659	3.96	***

Mean dependent var	2.323	SD dependent var	4.553
R-squared	0.983	Number of obs	2918
F-test	15715.535	Prob > F	0.000

**** p<.01, ** p<.05, * p<.1*

Heckman correction for all outcomes

ln_Q	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln_households	.846	.125	6.76	0	.601	1.091	***
ln_wage_2016	-1.344	.442	-3.04	.002	-2.21	-.479	***
ln_old_2016	.029	.246	0.12	.907	-.454	.511	
ln_young_2016	.422	.254	1.66	.096	-.075	.919	*
ln_productiv_2016	-.408	.446	-0.92	.36	-1.281	.465	
ln_unem_2016	-.05	.12	-0.41	.679	-.285	.186	
number of firms							
2	.578	.16	3.60	0	.263	.892	***
3	.845	.257	3.28	.001	.341	1.349	***
4	1.275	.367	3.47	.001	.556	1.994	***
5	1.288	.254	5.07	0	.79	1.786	***
Constant	14.84	3.405	4.36	0	8.167	21.513	***
ln_households	1.168	.096	12.19	0	.98	1.356	***
ln_wage_2016	.232	.421	0.55	.581	-.592	1.057	
ln_old_2016	.51	.196	2.60	.009	.125	.895	***
ln_young_2016	-.019	.188	-0.10	.918	-.388	.35	
ln_productiv_2016	.051	.354	0.14	.886	-.644	.745	
ln_unem_2016	.192	.099	1.93	.053	-.003	.386	*
ln_wage_pharmacists	-.529	.305	-1.73	.083	-1.126	.069	*
Constant	-6.634	3.699	-1.79	.073	-13.884	.616	*
athrho	.009	.066	0.13	.893	-.121	.139	
lnsigma	.151	.028	5.30	0	.095	.206	***
Mean dependent var		0.405	SD dependent var			1.043	
Number of obs		2918	Chi-square			632.107	
ρ = 0.008							
σ = 1.1625							
λ = 0.0103							

**** p<.01, ** p<.05, * p<.1*

Note, that values for $\rho, \sigma\ and\ \lambda$ are presented here as G transformed parameters.

Linear regression for OTC

ln_quantity_OTC	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln_households	.06	.039	1.57	.117	-.015	.136	
ln_wage_2016	-.304	.179	-1.70	.089	-.654	.047	*
ln_old_2016	.049	.065	0.75	.455	-.079	.176	
ln_young_2016	.025	.057	0.45	.654	-.086	.136	
ln_productiv_2016	-.001	.105	-0.01	.993	-.208	.206	
ln_unem_2016	.011	.044	0.25	.804	-.075	.096	
number of firms							
1	8.385	.056	151.07	0	8.276	8.494	***
2	9.945	.117	85.16	0	9.716	10.174	***
3	10.686	.187	57.06	0	10.319	11.053	***
4	11.219	.266	42.21	0	10.698	11.741	***
5	12.196	.123	98.91	0	11.954	12.438	***
Constant	1.449	1.281	1.13	.258	-1.062	3.96	

Mean dependent var	2.033	SD dependent var	4.081
R-squared	0.952	Number of obs	2918
F-test	5284.710	Prob > F	0.000

**** p<.01, ** p<.05, * p<.1*

Heckman selection model for OTC

ln_quantity_OTC	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln_households	.857	.199	4.32	0	.468	1.246	***
ln_wage_2016	-.984	.695	-1.42	.157	-2.345	.378	
ln_old_2016	-.216	.387	-0.56	.577	-.974	.542	
ln_young_2016	.627	.399	1.57	.116	-.154	1.409	
ln_productiv_2016	-.166	.7	-0.24	.812	-1.539	1.206	
ln_unem_2016	-.006	.189	-0.03	.973	-.377	.364	
number of firms							
2	.984	.252	3.90	0	.49	1.479	***
3	1.248	.405	3.09	.002	.456	2.041	***
4	1.497	.577	2.59	.009	.366	2.628	***
5	1.472	.402	3.66	0	.684	2.26	***
Constant	9.127	5.353	1.71	.088	-1.365	19.619	*
ln_households	1.169	.096	12.20	0	.981	1.356	***
ln_wage_2016	.236	.42	0.56	.574	-.587	1.059	
ln_old_2016	.513	.197	2.61	.009	.127	.898	***
ln_young_2016	-.018	.188	-0.10	.924	-.387	.351	
ln_productiv_2016	.047	.355	0.13	.895	-.648	.742	
ln_unem_2016	.192	.099	1.93	.053	-.003	.386	*
ln_wage_pharmacists	-.538	.306	-1.76	.079	-1.139	.062	*
Constant	-6.562	3.705	-1.77	.077	-13.823	.7	*
athrho	.024	.072	0.34	.733	-.116	.165	
lnsigma	.603	.028	21.22	0	.547	.659	***

Mean dependent var	0.405	SD dependent var	1.043
Number of obs	2918	Chi-square	357.781
$\rho = 0.0244$			
$\sigma = 1.8276$			
$\lambda = 0.0447$			

**** p<.01, ** p<.05, * p<.1*

Note, that values for $\rho, \sigma\ and\ \lambda$ are presented here as G transformed parameters.

Linear regression for PM

ln_medication	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln_households	.095	.027	3.58	0	.043	.147	***
ln_wage_2016	-.458	.123	-3.74	0	-.699	-.218	***
ln_old_2016	.047	.045	1.05	.293	-.041	.135	
ln_young_2016	-.013	.039	-0.32	.748	-.089	.064	
ln_productiv_2016	-.016	.072	-0.23	.821	-.158	.126	
ln_unem_2016	-.004	.03	-0.14	.89	-.063	.055	
number of firms							
1	9.567	.038	250.93	0	9.493	9.642	***
2	10.6	.08	132.14	0	10.443	10.758	***
3	11.167	.129	86.81	0	10.915	11.419	***
4	11.854	.183	64.93	0	11.496	12.212	***
5	12.654	.085	149.39	0	12.488	12.82	***
Constant	2.635	.88	3.00	.003	.91	4.36	***

Mean dependent var	2.230	SD dependent var	4.377
R-squared	0.980	Number of obs	2918
F-test	13265.193	Prob > F	0.000

**** p<.01, ** p<.05, * p<.1*

Heckman selection model for PM

ln_medication	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
ln_households	.823	.133	6.20	0	.563	1.083	***
ln_wage_2016	-1.444	.467	-3.09	.002	-2.359	-.528	***
ln_old_2016	-.044	.26	-0.17	.867	-.553	.466	
ln_young_2016	.181	.268	0.68	.499	-.344	.707	
ln_productiv_2016	-.107	.471	-0.23	.821	-1.03	.817	
ln_unem_2016	-.004	.127	-0.03	.974	-.253	.245	
number of firms							
2	.542	.17	3.20	.001	.21	.874	***
3	.734	.272	2.70	.007	.201	1.267	***
4	1.213	.388	3.13	.002	.453	1.974	***
5	1.209	.269	4.49	0	.681	1.736	***
Constant	14.779	3.6	4.11	0	7.724	21.834	***
ln_households	1.169	.096	12.19	0	.981	1.357	***
ln_wage_2016	.22	.42	0.52	.601	-.604	1.044	
ln_old_2016	.51	.196	2.59	.009	.124	.895	***
ln_young_2016	-.019	.188	-0.10	.92	-.388	.35	
ln_productiv_2016	.051	.354	0.14	.886	-.644	.745	
ln_unem_2016	.19	.099	1.92	.055	-.004	.385	*
ln_wage_pharmacists	-.521	.304	-1.71	.087	-1.118	.075	*
Constant	-6.62	3.699	-1.79	.074	-13.871	.63	*
athrho	-.016	.068	-0.24	.811	-.149	.117	
lnsigma	.206	.028	7.26	0	.151	.262	***

Mean dependent var	0.405	SD dependent var	1.043
Number of obs	2918	Chi-square	514.874

ρ =-.0162564

σ = 1.229025

λ = -.0199795

**** p<.01, ** p<.05, * p<.1*

Note, that values for $\rho, \sigma\ and\ \lambda$ are presented here as G transformed parameters.

Some code lines in Stata:

```
*clear
*use "C:\Users\ ... .dta"
//   ***collapse dataset to CITIES***
          //cities with N_pharmycy = 0 == 0
          *gen negative_city = 0
          *replace negative_city = -1 if city == .
          //collapse dataset - summary by cities
          *collapse (count) City_code (mean)
revenue_OTC_FINSTAT num_employee wages revenue_pres workday
nonstop days_open hospital employee_cat rural
wage_technicians density_2016 pop_2016 unem_2016 wage_2016
household_size pension_2017 elderly children population_grid
pharmacies_same_grid total_pharmacies_around district_code
region_code pct_slovak pct_roma pct_hungarian educ_prim
educ_sec_grad educ_sec educ_univ educ_none median_income
density_1996 density_1997 density_1998 density_1999
density_2000 density_2001 density_2002 density_2003
density_2004 density_2005 density_2006 density_2007
density_2008 density_2009 density_2010 density_2011
density_2012 density_2013 density_2014 density_2015 pop_1996
pop_1997 pop_1998 pop_1999 pop_2000 pop_2001 pop_2002
pop_2003 pop_2004 pop_2005 pop_2006 pop_2007 pop_2008
pop_2009 pop_2010 pop_2011 pop_2012 pop_2013 pop_2014
pop_2015 old_1996 old_1997 old_1998 old_1999 old_2000
old_2001 old_2002 old_2003 old_2004 old_2005 old_2006
old_2007 old_2008 old_2009 old_2010 old_2011 old_2012
old_2013 old_2014 old_2015 old_2016 young_1996 young_1997
young_1998 young_1999 young_2000 young_2001 young_2002
young_2003 young_2004 young_2005 young_2006 young_2007
young_2008 young_2009 young_2010 young_2011 young_2012
young_2013 young_2014 young_2015 young_2016 productiv_1996
productiv_1997 productiv_1998 productiv_1999 productiv_2000
productiv_2001 productiv_2002 productiv_2003 productiv_2004
productiv_2005 productiv_2006 productiv_2007 productiv_2008
productiv_2009 productiv_2010 productiv_2011 productiv_2012
productiv_2013 productiv_2014 productiv_2015 productiv_2016
roma_2001 roma_2002 roma_2003 roma_2004 roma_2005 roma_2006
roma_2007 roma_2008 roma_2009 roma_2010 roma_2011 roma_2012
roma_2013 roma_2014 roma_2015 roma_2016 hungar_2001
hungar_2002 hungar_2003 hungar_2004 hungar_2005 hungar_2006
hungar_2007 hungar_2008 hungar_2009 hungar_2010 hungar_2011
hungar_2012 hungar_2013 hungar_2014 hungar_2015 hungar_2016
unem_2001 unem_2002 unem_2003 unem_2004 unem_2005 unem_2006
unem_2007 unem_2008 unem_2009 unem_2010 unem_2011 unem_2012
unem_2013 unem_2014 unem_2015 unem_2017 wage_2009 wage_2010
wage_2011 wage_2012 wage_2013 wage_2014 wage_2015 households
```

```
num_firms_registered men_2016_00_04 men_2016_05_09
men_2016_10_14 men_2016_15_19 men_2016_20_24 men_2016_25_29
men_2016_30_34 men_2016_35_39 men_2016_40_44 men_2016_45_49
men_2016_50_54 men_2016_55_59 men_2016_60_64 men_2016_65_69
men_2016_70_74 men_2016_75_79 men_2016_80_84 men_2016_85_89
men_2016_90_94 men_2016_95_99 men_2016_100_104
men_2016_105_109 men_2016_110_114 women_2016_00_04
women_2016_05_09 women_2016_10_14 women_2016_15_19
women_2016_20_24 women_2016_25_29 women_2016_30_34
women_2016_35_39 women_2016_40_44 women_2016_45_49
women_2016_50_54 women_2016_55_59 women_2016_60_64
women_2016_65_69 women_2016_70_74 women_2016_75_79
women_2016_80_84 women_2016_85_89 women_2016_90_94
women_2016_95_99 women_2016_100_104 women_2016_105_109
women_2016_110_114 wage_pharmacists negative_city (sum)
quantity_OTC revenue_OTC quantity_medication
revenue_medication, by (City)
          //set actual city count ~ N_pharmacy
               *gen N_pharmacy = City_code + negative_city
               *drop City_code negative_city
               *label variable N_pharmacy "number of firms =
pharmacies"
               *drop if N_pharmacy >5
          //Vars
          *gen Q = quantity_OTC+quantity_medication
          *gen ln_Q = log(quantity_OTC+quantity_medication)
          *replace ln_Q = 0 if ln_Q == .
          *label variable ln_Q "log of
quantity_OTC+quantity_medication"
          //log of population
          *gen ln_pop_2016 = log(pop_2016)
          *label variable ln_pop_2016 "log of pop_2016"
          //log vars
          *gen ln_wage_2016 = log(wage_2016)
          *label variable ln_wage_2016 "Dδ & Vα"
          *gen ln_density_2016 = log(density_2016)
          *label variable ln_density_2016 "Dδ"
          *gen ln_old_2016 =log(old_2016)
          *label variable ln_old_2016 "Dδ"
          *gen ln_young_2016 = log(young_2016)
          *label variable ln_young_2016 "Dδ"
          *gen ln_productiv_2016 =log(productiv_2016)
          *label variable ln_productiv_2016 "Dδ & Vα"
          *gen ln_pct_slovak = log(pct_slovak)
          *label variable ln_pct_slovak "Dδ"
          *gen ln_pct_roma = log(pct_roma)
          *label variable ln_pct_roma "Dδ"
          *gen ln_pct_hungarian = log(pct_hungarian)
```

```
          *label variable ln_pct_hungarian "Dδ"
          *gen ln_roma_2016 = log(roma_2016)
          *label variable ln_roma_2016 "Dδ"
          *gen ln_hungar_2016 = log(hungar_2016)
          *label variable ln_hungar_2016 "Dδ"
          *gen ln_unem_2016 = log(unem_2016)
          *label variable ln_unem_2016 "Dδ & Vα"
          *gen ln_household_size = log(household_size)
          *label variable ln_household_size "Sλ"
          *gen ln_households = log(households)
          *label variable ln_households "Sλ"
          *gen ln_pension_2017 = log(pension_2017)
          *label variable ln_pension_2017 "Dδ & Vα"
          //N_pharmacy vars
          *gen ln_wages = log(wages)
          *gen ln_workday = log(workday)
          //nonstop
          *gen ln_day_open = log(days_open)
          //hospital
          //rural
          *gen ln_quantity_OTC = log(quantity_OTC)
          *gen ln_revenue_OTC  = log(revenue_OTC)
          *gen ln_wage_pharmacists = log(wage_pharmacists)
          *label variable ln_wage_pharmacists "Wγ"
          *gen ln_wage_technicians = log(wage_technicians)
          *gen ln_revenue_medication =
log(revenue_medication)
          *gen ln_elderly = log(elderly)
          *gen ln_children = log(children)
          *gen ln_educ_none = log(educ_none)
          *gen ln_educ_prim = log(educ_prim)
          *gen ln_educ_sec = log(educ_sec)
          *gen ln_educ_sec_grad = log(educ_sec_grad)
          *gen ln_educ_univ = log(educ_univ)
          *gen ln_median_income = log(median_income)
/// DP
*keep N_pharmacy Q ln_pop_2016 ln_households ln_wage_2016
ln_density_2016 ln_old_2016 ln_young_2016 ln_productiv_2016
ln_roma_2016 ln_hungar_2016 ln_unem_2016 ln_pension_2017
ln_wage_pharmacists
*gen ln_Q = log(Q)
*replace N_pharmacy = 5 if N_pharmacy >5
*sum Q households wage_2016 old_2016 young_2016
productiv_2016 unem_2016 wage_pharmacists
*asdoc sum Q households wage_2016 old_2016 young_2016
productiv_2016 unem_2016 wage_pharmacists
```

```
*asdoc oprobit N_pharmacy ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists
*asdoc regress ln_Q ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016 i.N_pharmacy
*heckman ln_Q ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016 i.N_pharmacy,
select(N_pharmacy = ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
*replace ln_quantity_OTC = 0 if ln_quantity_OTC == .
*asdoc regress ln_quantity_OTC ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
i.N_pharmacy
*asdoc heckman ln_quantity_OTC ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
i.N_pharmacy, select(N_pharmacy = ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
*replace ln_medication = 0 if ln_medication == .
*asdoc regress ln_medication ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
i.N_pharmacy
*asdoc heckman ln_medication ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
i.N_pharmacy, select(N_pharmacy = ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
//ETR
*oprobit N_pharmacy ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists
**Save the matrix (scalars) by "mat" comand of cut values
(constants) - 8 cuts = 9-1 as start form N=1 to N=10
*mat delta_1 = _b[/cut1]
*matlist delta_1
*mat delta_2 = _b[/cut2]
*mat delta_3 = _b[/cut3]
*mat delta_4 = _b[/cut4]
*mat delta_5 = _b[/cut5]
//Save lambda as a coefficient for the market size
*mat lambda = _b[ln_households]
*matlist lambda
//Save all coefficients into the B matrix
*mat b = e(b)
*matlist b
//Save coefficients for other control variales (vector X & W,
for now n(b)=8)
```

```
*mat beta = b[1,2..6]
*matlist beta
*matlist beta'
//Make average of control variables
*mean ln_wage_2016 ln_old_2016 ln_young_2016
ln_productiv_2016 ln_unem_2016
//Save means of control variables as vector "means_x"
*mat means_x = e(b)
*matlist means_x
*matlist means_x'
//Vector "means_x" multiplied by transpsed vector (') "beta"
(its coefficients)
*mat means_xb=beta*means_x'
*matlist means_xb
//////////////////////////////////////////////////////////////
/////////
**counting ETs
//Count the ET for firm 1 "S1"
*mat S1 = -means_xb + delta_1
*matlist S1
//divide by lambda as a coefficient for the market size
*mata : st_matrix("S1", st_matrix("S1") :/
st_matrix("lambda"))
*matlist S1
//from log pops to real pops
*mata : st_matrix("S1", exp(st_matrix("S1")))
//actual ET for the first firm
*matlist S1
//Count the ET for firm 2 "S2"
*mat S2 = -means_xb + delta_2
*matlist S2
//divide by lambda as a coefficient for the market size
*mata : st_matrix("S2", st_matrix("S2") :/
st_matrix("lambda"))
*matlist S2
//from log pops to real pops
*mata : st_matrix("S2", exp(st_matrix("S2")))
//actual ET for the second firm
*matlist S2
//outcome for ET3 - third firm
*mat S3 =  -means_xb + delta_3
*mata : st_matrix("S3", st_matrix("S3") :/
st_matrix("lambda"))
*matlist S3
*mata : st_matrix("S3", exp(st_matrix("S3")))
*matlist S3
//outcome ET4
*mat S4 =  -means_xb + delta_4
```

```
*mata : st_matrix("S4", st_matrix("S4") :/
st_matrix("lambda"))
*matlist S4
*mata : st_matrix("S4", exp(st_matrix("S4")))
*matlist S4
//outcome ET4
*mat S5 =  -means_xb + delta_5
*mata : st_matrix("S5", st_matrix("S5") :/
st_matrix("lambda"))
*matlist S5
*mata : st_matrix("S5", exp(st_matrix("S5")))
*matlist S5
//ET
*matlist S1
*matlist S2
*matlist S3
*matlist S4
*matlist S5
//ET for only one firm, ET per one firm
*mat s1 = S1/1
*mat s2 = S2/2
*mat s3 = S3/3
*mat s4 = S4/4
*mat s5 = S5/4
//List of ETs
*matlist s1
*matlist s2
*matlist s3
*matlist s4
*matlist s5
//////////////////////////////////////////////////////////////
/////////
// Create a single matrix containing all expected values
*matrix all_expected_values = s1, s2, s3, s4, s5
*matlist all_expected_values
//////////////////////////////////////////////////////////////
/////////

// Calculate ETRs between firms
*mata : st_matrix("ETR21_nosel", st_matrix("s2") :/
st_matrix("s1"))
*mata : st_matrix("ETR32_nosel", st_matrix("s3") :/
st_matrix("s2"))
*mata : st_matrix("ETR43_nosel", st_matrix("s4") :/
st_matrix("s3"))
*mata : st_matrix("ETR54_nosel", st_matrix("s5") :/
st_matrix("s4"))
//actual ETRs
```

```
*matlist ETR21_nosel
*matlist ETR32_nosel
*matlist ETR43_nosel
*matlist ETR54_nosel
// Calculate pseudo ETRs
*mata : st_matrix("p_ETR21_nosel", st_matrix("s2") :/
st_matrix("s1"))
*mata : st_matrix("p_ETR32_nosel", st_matrix("s3") :/
st_matrix("s1"))
*mata : st_matrix("p_ETR43_nosel", st_matrix("s4") :/
st_matrix("s1"))
*mata : st_matrix("p_ETR54_nosel", st_matrix("s5") :/
st_matrix("s1"))
//Relative ETRs
*matlist p_ETR21_nosel
*matlist p_ETR32_nosel
*matlist p_ETR43_nosel
*matlist p_ETR54_nosel
//Demand
*heckman ln_Q ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016 i.N_pharmacy,
select(N_pharmacy = ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
*matrix d_ = e(b)
*matlist d_
*mat demand_N2 = d_[1,8]
*matlist demand_N2
*mat demand_N3 = d_[1,9]
*matlist demand_N3
*mat demand_N4 = d_[1,10]
*matlist demand_N4
*mat demand_N5 = d_[1,11]
*matlist demand_N5
*mata :
st_matrix("sel_e_demand_N2",exp(st_matrix("demand_N2")))
*matlist sel_e_demand_N2
*mata :
st_matrix("sel_e_demand_N3",exp(st_matrix("demand_N3")))
*matlist sel_e_demand_N3
*mata :
st_matrix("sel_e_demand_N4",exp(st_matrix("demand_N4")))
*matlist sel_e_demand_N4
*mata :
st_matrix("sel_e_demand_N5",exp(st_matrix("demand_N5")))
*matlist sel_e_demand_N5
*heckman ln_quantity_OTC ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
```

```
i.N_pharmacy, select(N_pharmacy = ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
*matrix d_ = e(b)
*matlist d_
*mat demand_N2 = d_[1,8]
*matlist demand_N2
*mat demand_N3 = d_[1,9]
*matlist demand_N3
*mat demand_N4 = d_[1,10]
*matlist demand_N4
*mat demand_N5 = d_[1,11]
*matlist demand_N5
*mata :
st_matrix("sel_e_demand_N2",exp(st_matrix("demand_N2")))
*matlist sel_e_demand_N2
*mata :
st_matrix("sel_e_demand_N3",exp(st_matrix("demand_N3")))
*matlist sel_e_demand_N3
*mata :
st_matrix("sel_e_demand_N4",exp(st_matrix("demand_N4")))
*matlist sel_e_demand_N4
*mata :
st_matrix("sel_e_demand_N5",exp(st_matrix("demand_N5")))
*matlist sel_e_demand_N5
heckman ln_Q ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016 i.N_pharmacy,
select(N_pharmacy = ln_households ln_wage_2016 ln_old_2016
ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
heckman ln_quantity_OTC ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
i.N_pharmacy, select(N_pharmacy = ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
*heckman ln_quantity_medication ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
i.N_pharmacy, select(N_pharmacy = ln_households ln_wage_2016
ln_old_2016 ln_young_2016 ln_productiv_2016 ln_unem_2016
ln_wage_pharmacists)
*matrix d_ = e(b)
*matlist d_
*mat demand_N2 = d_[1,8]
*matlist demand_N2
*mat demand_N3 = d_[1,9]
*matlist demand_N3
*mat demand_N4 = d_[1,10]
*matlist demand_N4
```

```
*mat demand_N5 = d_[1,11]
*matlist demand_N5
*mata :
st_matrix("sel_e_demand_N2",exp(st_matrix("demand_N2")))
*matlist sel_e_demand_N2
*mata :
st_matrix("sel_e_demand_N3",exp(st_matrix("demand_N3")))
*matlist sel_e_demand_N3
*mata :
st_matrix("sel_e_demand_N4",exp(st_matrix("demand_N4")))
*matlist sel_e_demand_N4
*mata :
st_matrix("sel_e_demand_N5",exp(st_matrix("demand_N5")))
*matlist sel_e_demand_N5
```