Instrument-free structural estimation of the nursing home market under entry-exit actions

Ken Inosaki*

November 15, 2024

Abstract

This research introduces a novel approach for estimating the structural parameters of demand, cost, and entry costs in a differentiated products model where product characteristics and input cost data are not observed for non-entrants. Traditional methods for entry game estimation rely on the product characteristics that are used as instruments to be observable for both entrants and non-entrants — a scenario that is uncommon in practice. I first provide an extension of the standard identification strategy that does not require such observability condition, but also demonstrate based on identification analysis as well as Monte-Carlo study that such an approach requires impractically large sample size.

To overcome this limitation, I use the instrument-free methods proposed by Byrne et al. (2022) and Imai et al. (2024), which allow estimation of the demand and cost function by addressing the endogeneity of price using entrants' cost data. Building upon this foundation, I extend their framework to incorporate entry-exit decisions. My findings indicate that using both demand and cost data offers a more practical and effective estimation approach. I propose a data-augmented Markov Chain Monte Carlo (MCMC) estimation method and demonstrate through Monte Carlo simulations that this approach yields consistent estimates.

Furthermore, I apply the estimation techniques developed in this research to estimate the structural parameters of the Wisconsin nursing home market and discuss the social welfare implications of the Certificate of Need (CON) law. Counterfactual simulations reveal that abolishing the CON law would increase consumer and producer surplus by $868 million and $165 million, respectively, while government spending would rise by $700 million. I also estimate important market structures, such as labor/capital elasticities, entry costs, and the difference in the distribution of service quality between entrants and non-entrants.

*Department of Economics, University of Washington at Seattle. email: inosaki@uw.edu

1 Introduction

I study the differentiated products oligopoly model in which firms decide whether to enter the market in period 1 and engage in a price competition in period 2. The frameworks of interest for this project are the logit and random coefficient logit models of Berry (1994) and Berry et al. (1995) (hereafter, BLP) which are used to estimate demand and cost parameters of firms in a differentiated products oligopoly market. These models incorporate unobserved heterogeneity in product quality. Researchers use instrumental variables to deal with the endogeneity of prices caused by such product quality. As Berry and Haile (2014) and others point out, as long as some instruments are available, demand functions can be identified using market-level data. Popular instruments include cost shifters such as market input prices, product characteristics of other products in a market ("BLP instruments"), and the price of a given product in other markets ("Hausman instruments"). However, these instruments become invalid when we explicitly consider firms' entry and exit decisions. Note that the demand function estimation methods based on Berry (1994) and Berry et al. (1995)'s models do not consider firms' entry-exit decisions. As I will show later, the well-known instruments for price become invalid if firms make entry decisions due to the endogeneity of entry.

There are studies, such as Ciliberto and Tamer (2009), Ciliberto et al. (2021), Aguirregabiria et al. (2024), and others, that structurally estimate parameters in oligopoly markets where firms make endogenous entry decisions. All of these studies analyze the airline industry. The primary appeal of the airline industry for researchers is that, at each airport, the potential entrants are existing airlines whose characteristics are observable. This ideal feature allows researchers to identify the conditional probability of firm entry based on observable characteristics. However, it is rare to have data on potential entrants in other industries; typically, researchers only have information on firms that actually entered the market.

In this project, I identify and estimate the parameters of demand and cost functions in a model of a differentiated products oligopoly entry game when data on non-entrant firms are not available. In such cases, even though researchers may have valid instruments that are orthogonal to the error term in the population—including both entrants and non-entrants—they can only construct moments based on data from entrants. These moment conditions may not reflect the orthogonality conditions in the overall population. To correct for the selection bias, researchers typically use entry probabilities or similar statistics. However, estimating the entry probability as a function of exogenous variables requires data on those variables for non-entrants, which may not be available in most industries. Instead, I assume in my model that researchers do not have any data on non-entrant firms but can obtain data on demand and cost of the entrants (total cost and each input costs, but not the marginal cost). In many industries under government oversight, such as nursing homes, banking, etc., incumbent firms have to report to the oversight authorities some data on cost.

This research contributes to the extensive literature on structural estimation of static entry games. Beginning with foundational works by Tobin (1958), Amemiya (1973), and Heckman (1976), economists have extensively examined the impact of selection effects on outcomes of interest. Bresnahan et al. (1987), Bresnahan and Reiss (1990), and Bresnahan and Reiss (1991) pioneered the estimation of entry game models, setting the stage for subsequent research in this area. Studies such as Reiss and Spiller (1989), Berry (1992), Mazzeo (2002), Seim (2006), Jia (2008), and Ciliberto and Tamer (2009) delve into entry models where the presence of competing firms influences profits in a (log) linear framework. In contrast, Wollmann (2018), Ciliberto et al. (2021), and Aguirregabiria et al. (2024) develop entry models based on demand structures from Berry (1994), where the strategic entry effect is embedded within a structural model. Related research, such as Draganska et al. (2009), Sweeting (2013), Eizenberg (2014), and Li et al. (2022), explores models where firms first select their product characteristics in an initial period and subsequently compete with other entrants. Similarly, Ho (2009), Kuehn (2018), Park (2020), Bontemps et al. (2023), and Yuan and Jia (2024) construct network models where firms decide on their network configurations in the first period and then compete with rivals in the second period. The model developed in this article diverges from prior models in that it does not require observing the product characteristics of non-entrants, unlike previous models that depend on such information for identification purposes.

To identify the structural parameters, I use the methodologies developed in Byrne et al. (2022) and Imai et al. (2024). In particular, Imai et al. (2024) not only remove the measurement error but also identify and estimate the price coefficients and the output coefficient of the cost function without instruments. They use the marginal cost function to proxy for the unobservable cost shock and utilize MR = MC equation derived from the profit maximization problem so that the marginal cost can be expressed as a function of observable demand variables and demand parameters. Demand and cost parameters of this modified cost function are chosen to maximize the fit to the cost data. Since the methodology is instrument-free, researchers can estimate those coefficients even if they only use the data on entrants for estimation, in which case the orthogonality conditions based on instruments do not hold. The shortcoming of their approach is that they cannot identify some of the demand and cost functions. For example, their approach cannot identify the coefficient of the product characteristics which usually enter in the logit utility.After adopting the methodology outlined by Imai et al. (2024) to identify a subset of the demand and cost function parameters, I demonstrate how this enables the identification of profits. This, in turn, enables the determination of their conditional probability of entering the market. By conditioning on this entry probability, we can remove the bias introduced by the endogeneity of market entry and estimate the remaining structural parameters.

I show the identification and propose an estimation procedure based on data-augmented MCMC algorithm. The results show that the proposed estimator can accurately estimate the structural parameters when we have both demand and cost data on new entrants. Moreover, I show that the structural parameters can be estimated even when the researcher does not have the cost data, only when the variance of the exogenous variables is small.

I use the model I developed in this article to estimate the market structure of the Wisconsin skilled nursing facility(SNF) market and assess the Certificate of Need (CON) law's impact on social welfare.

With the increasing elderly population, Wisconsin has seen its state expenditures for nursing homes rise significantly. To curb these escalating costs, the state implemented the CON law. This legislation restricts the construction of new nursing facilities and the expansion of existing ones.[1] Although the CON law may have reduced government spending, a large body of literature has studied its negative effects due to excess demand in the industry.

First, Gruenberg and Willemain (1982), Gertler (1989), and Gertler (1992) use both economic theory and empirical studies to verify the relationship between excess demand and patients' access to nursing homes. Limited supply led Medicaid patients to be rationed out, since their reimbursement rate was significantly lower than that of private-pay patients (PPPs). This is problematic for at least a couple of reasons. Firstly, low-income potential patients who need professional assistance may not be able to get any treatment. Secondly, as Ettner (1993) pointed out, the average hospital cost per patient-day for a semiprivate hospital room was \$465 in 1985, while the average nursing homes' private patient price was only \$1,456 per month. Since 40% of nursing home residents come directly from hospitals, they note that supply constraints may create inefficient government spending. Kotschy and Bloom (2022) examined data from 30 developed countries and found that difficulty in accessing nursing homes is a common issue globally.

[1] In fact, according to my data, none of the nursing homes that were present from 1998 to 2002 changed the number of beds during that time.

Second, Nyman (1985), Nyman (1988a), Nyman (1988b), Harrington et al. (2000), Horn et al. (2005), Lin (2015), Lu et al. (2021), and Kunz et al. (2024) consider various measures of nursing home quality and discuss its relationship with excess demand, most of them reporting a negative relationship between the two variables. Nyman (1988a) points out that nursing homes' low quality was a well-known issue, as officially reported in the senate report in 1994[2]. Due to low Medicaid reimbursement and flat rates, reducing costs rather than engaging in quality competition provided greater benefits for nursing homes. Gupta et al. (2024) also discuss the relationship between health condition and private equity and finds that private equity ownership increases mortality rate by 10%. Third, Norton (1992), Cohen and Spector (1996) , Grabowski (2001), and Grabowski and Angelelli (2004), focus on Medicaid reimbursement rates and their implications in relation to market outcomes. Especially, Grabowski (2001) and Grabowski and Angelelli (2004) find a positive effect of an increase in Medicaid reimbursement rate on health outcomes, although the CON may mitigate this effect.[3] Fourth, Nyman (1994) explores the relationship between price and excess demand, which they find a positive relationship between the two variables. Bardey and Siciliani (2021) uses two-sided economic model and finds that the profits and wages for nurses become lower when the prices are regulated, and Heger et al. (2022) uses exogenous variation in Swiss care price regulation and finds that higher prices leads to higher staffing ratio. Yang et al. (2022) offer a literature review regarding pricing behavior and concentration in the industry.

However, as pointed out in Ching et al. (2015), most of the empirical strategies in the existing literature use reduced-form estimation methods, which have several limitations. First, they cannot estimate social welfare. Second, their estimation strategies often suffer from endogeneity issues. For example, Gertler (1989) analyzes the effect of excess demand on quality by running regressions that include measures of excess demand and the Herfindahl-Hirschman Index (HHI) in the right hand side to control for market structure. However, as Miller et al. (2022) points out, both excess demand and HHI are market outcomes influenced by observed and unobserved shocks. As a result, it is impossible to establish causal relationship between the two endogenous variables. Analysis of effect of excess demand on market structure requires a structural approach.

Recent papers have begun using structural approaches to analyze the effect of excess demand and the market structure. Ching et al. (2015) develop a static oligopoly model based on Berry

[2]The exact same senate report cited in their paper was not found on internet, so I leave below the link to its introductory report: https://archive.org/details/nursinghomecarei00unit/page/n1/mode/2up

[3]Regarding the Medicare program, He et al. (2020) finds a positive relationship between Medicare reimbursement rate and staffing level.

et al. (1995) to estimate private-patient demand. They use the estimated parameters to quantify the rationing effect of the CON law. However, since they only model the demand side, they cannot estimate the producer-side welfare, which is also essential in discussing the CON law's effect. Moreover, their static model cannot fully analyse how profit competition between firms promotes quality competition. Hackmann (2019) extends Ching et al. (2015)'s study by endogenizing nursing homes' quality and explicitly modeling the cost side. In addition to the static models shown above, there are several papers that use dynamic models to study the market structure. Gowrisankaran and Town (1997) was one of the pioneers that used a full dynamic model that levaraged Ericson and Pakes (1995) model to estimate demand, cost, entry/exit cost, income elasticity, and other parameters of the hospital market. This paper shows groundbreaking results and implications that has a lot of common things with the nursing home industry, but it is not possible to discuss the level of rationing with this economic model since they do not explicitly incorporate the rationing behavior in their model. Lin (2015) establishes a dynamic model following Ericson and Pakes (1995), where nursing homes make dynamic decisions about their treatment quality. They use panel data to identify the competition effect of profit and entry cost. Nonetheless, their model is not suitable for analyzing the CON law as they do not model consumer choices and the cost of production, making it impossible to quantify the social effect of the CON law or the rationing effect. Grant et al. (2022) uses Germany nursing homes' market data and develops an entry model to study the competition of for-profit and non-profit nursing homes. Since they simplify the profit function in order to focus on dynamic entry game behavior, and Germany does not have laws similar to the CON law, measuring the effect of the CON using this model is not suitable. Overall, there are no research papers that thoroughly estimate the CON law's effect and quality competition effect incorporating entry-exit models.

The research most closely related to mine is Hackmann (2019), who developed a model similar to that of Ching et al. (2015). In their study, they endogenize nursing home quality, i.e. the number of nurses per patient, and use a constant marginal cost function, estimating the model using instrumental variables (IV). However, their model has several limitations. First, their identification strategy using IV may be biased. They use Medicaid patients' reimbursement rate as an instrument for the demand shock, assuming that the cost variation is orthogonal to unobserved preference shocks in the given nursing home county. However, this assumption does not always hold, especially if the cost shock is serially correlated or correlated with the cost structure, which cannot be verified. Second, they do not model the entry behavior of nursing homes. Although the CON law restricts the construction of new nursing homes, the data in

Figure 1 show that the number of facilities has fluctuated significantly each year. Therefore, as discussed section 4, the structural parameters can be biased when the entry behavior of firms is not taken into account. Third, they use a constant marginal cost function, which lacks flexibility due to the lack of cost data. As demonstrated in my research, nursing homes exhibit increasing returns to scale, which could bias their counterfactual simulations analyzing the impact of an increase in Medicaid patients' reimbursement rates on facility quality. Unlike Hackmann (2019), the model I use can estimate the cost function flexibly, as the estimation strategy does not require IV.

This paper uses the structural approach developed in section 4 to answer the following two major questions prevalent in Wisconsin's skilled nursing home facility market:

1. What is the quantitative effect of the CON law on social welfare and rationing?

2. Does competition improve the quality of treatment?

Most of the previous studies mentioned so far have analyzed various types of nursing facilities including skilled nursing facilities, nursing homes connected to hospitals, and home health agencies without making clear distinctions between them. In my research, however, instead of analyzing all types of nursing homes, I focus only on skilled nursing facility due to data limitation. According to California Department of Aging,[4] a typical resident of SNF is 'a person who is chronically ill or recuperating from an illness or surgery and needs regular nursing care and other health related services.' Unlike other types of nursing homes, a SNF is ' is a temporary residence for patients undergoing medically necessary rehabilitation treatment.'[5]

Following Ching et al. (2015), the model I use assumes that private-pay patients (PPP) do not face rationing, while Medicaid patients (MP) do face rationing. I use PPP's discrete choice outcomes to estimate the structural parameters and conduct counterfactual simulations of MP using the estimated structure.

This article also contributes to the literature on the estimation of nursing home cost functions. The literature has used a variety of methods to measure the cost structure of nursing homes. Vitaliano and Toren (1994), Hofler and Rungeling (1994), and Mutter et al. (2013) use the stochastic frontier approach to examine the inefficiency of nursing homes. Bekele and Holtmann (1987), Gertler and Waldman (1992), Filippini (2001), and Giorgio et al. (2016) use the translog cost function structure to estimate flexible cost functions. Dudzinski et al. (1998)

[4] https://www.aging.ca.gov/Care_Options/Skilled_Nursing_Facilities/

[5] https://www.hebrewseniorlife.org/blog/difference-between-nursing-homes-and-skilled-nursing-facilities

use the Hedonic approach to find nursing home's returns to scale, and Christensen (2004) use quantile regression with a translog form to discuss the heterogeneity in efficiency. Some other papers use reduced-form approaches; Knox et al. (2004) use regression models to discuss the heterogeneity of efficiencies across different types of nursing homes, and Murmann et al. (2023) use the propensity-score method to discuss the cost-effectiveness of transitional care in Canada.

However, to the best of my knowledge, none of them address the endogeneity of quantity.[6] Production quantity is determined by the firm's profit maximization problem. Therefore, firms with higher marginal costs (or higher cost shocks) produce fewer products. Consequently, there is likely a correlation between observed production quantity and unobserved heterogeneity in cost structures or cost shocks. This issue has been overlooked in the literature due to a lack of valid instrumental variables. In contrast, the estimation method used in this article circumvents this problem by canceling out the unobserved cost components from the estimation equations. I estimate the cost function derived from the Cobb-Douglas production function and discuss its implications.

This paper is also one of the few to provide estimates of entry costs. Most papers rely on the dynamic behaviour of firms to identify them using panel data. They find different estimation results regarding entry costs; Gowrisankaran and Town (1997) finds the entry cost is statistically insignificant to zero, while Lin (2015) and Grant et al. (2022) finds a large entry cost. My article identifies the entry cost in a static game from a novel perspective; it identifies the entry cost from firms' profit distribution. Entrant firms having large profit implies high entry cost in my model because otherwise it cannot explain other potential entrants not entering the market despite the market being profitable. I show that the entry cost is relatively small compared to the total cost, as the mean of the profit distribution of SNF is close to zero.

The estimation results of my model carry significant policy implications. Firstly, the labor and capital elasticities are 0.698 and 0.662, respectively, summing to 1.361. This result implies increasing returns to scale, a finding corroborated by previous research(e.g. Bekele and Holtmann (1987), Filippini (2001)). Such insights are crucial for policymaking, indicating that government initiatives to invest in SNFs would be more cost-effective if focused on expanding larger establishments rather than building many smaller ones. Secondly, the estimated daily entry cost for a nursing home is \$10.9, amounting to an annual cost of \$4,102. This figure is considerably lower compared to the total operational costs, which average \$11,704 daily, or approximately \$4.27

[6] Gertler and Waldman (1992) and Giorgio et al. (2016) address the issue of quality endogeneity using econometric theories and instrumental variables. However, the problem of endogeneity in quantity remains unaddressed.

million annually. This disparity likely stems from intense competition within the SNF sector, with the entry cost in my model identified from firms' profitability.

In conducting a series of counterfactual simulations using estimated demand, cost, and entry parameters, I first evaluated the impact of the Certificate of Need (CON) law on rationing effects and social welfare. Assuming the absence of the CON law, an additional 17,287 patients would have access to SNFs. Furthermore, both consumer and producer surplus would increase by $868 million and $165 million annually, respectively. Meanwhile, government expenditure would rise by $700 million, leading to an annual increase in social welfare of $333 million. This result supports the findings of Ching et al. (2015), which point out that current policy results in a net welfare loss. Additionally, I analyzed the distribution of quality among nursing homes, comparing current entrants to potential entrants. Utilizing Ching et al. (2015)'s measure of nursing home quality, defined as $\mathbf{x}\boldsymbol{\beta}+\xi$, I discovered that the average quality measure of entrants is 9.6% higher than that of potential entrants. This result indicates that entry-exit competition significantly improves quality, offering a potential solution to the problem of substandard quality in nursing homes.

This article is organized as follows. Section 2 briefly reviews the literature and present the entry model I estimate. Then, in Section 3, I describe the first stage identification strategy based on Imai et al. (2024). Section 4 discusses how I identify the remaining parameters of the market share and cost function, the distribution of observed and unobserved product characteristics and the cost shock, and the entry cost without data on non-entrants in a monopoly market. Section 5 discuss extends the monopoly model to a oligopoly model and logit model to the BLP model. Section 6 applies the proposed estimation method to Wisconsin's SNF market to estimate demand, cost, and entry parameters to analyze the effect of the CON law.

2 The two-period model of entry and price competition in a differentiated products oligopoly market.

I follow Ciliberto et al. (2021) and Aguirregabiria et al. (2024) by considering a two-period model where in period 1, potential entrants make entry decisions, and in period 2, only entrants in the market engage in price competition.

I first review the standard differentiated products demand model, then the cost function. I next discuss the profit maximizing price setting behavior of entrants, and finally, period 1 where firms make entry decisions.

2.1 Differentiated products discrete choice demand models based on logit market share

In this subsection, I describe the standard differentiated products model. For more details, see Berry (1994), Berry et al. (1995), Nevo (2001) and others. Most features of the model I discuss here are carried over to the next section where I explain my cost data-based identification strategy.

I use the notation following Byrne et al. (2022). Consumer i in market m gets the following utility from consuming one unit of product j:

$$u_{ijm} = \mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm} + \epsilon_{ijm},$$

where $\mathbf{x}_{jm}$ is a $1 \times K$ vector of observed product characteristics, p_{jm} is price, ξ_{jm} is the unobservable demand shock, and ϵ_{ijm} is an idiosyncratic taste shock. Let the demand parameter vector be $\boldsymbol{\theta}_d = \left[\alpha, \boldsymbol{\beta}'\right]'$. $\boldsymbol{\beta}$ is a $K \times 1$ vector.

There are $m = 1 \dots M$ markets that have market sizes Q_m.[7] Each market has $j = 0 \dots J_m$ products whose aggregate demand across individuals is

$$q_{jm} = s_{jm} Q_m,$$

where q_{jm} denotes output and s_{jm} denotes market share. In the case of the Berry (1994) logit demand model, ϵ_{ijm} are assumed to be i.i.d. across products, individuals and markets,and are type I extreme value distributed. Then, the aggregate market share for product j in market m is,

$$s_{jm}\left(\boldsymbol{\theta}_d\right) \equiv s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}_d\right) = \frac{\exp\left(\mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}\right)}{\sum_{k=0}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)}, \tag{1}$$

where $\mathbf{p}_m = [p_{0m}, p_{1m}, ..., p_{J_m m}]'$ is a $(J_m + 1) \times 1$ vector,

$$\mathbf{X}_m = \begin{bmatrix} \mathbf{x}_{0m} \\ \mathbf{x}_{1m} \\ \vdots \\ \mathbf{x}_{J_m m} \end{bmatrix}$$

is a $(J_m + 1) \times K$ matrix, $\boldsymbol{\xi}_m = [\xi_{0m}, \xi_{1m}, ..., \xi_{J_m m}]'$ is a $(J_m + 1) \times 1$ vector.

[7] With panel data the m index corresponds to a market-period.

Good $j = 0$ represents the outside option that corresponds to not buying any of the $j = 1, \ldots, J_m$ goods. Outside option's product characteristics, price, and demand shock are normalized to zero (i.e., $\mathbf{x}_{0m} = \mathbf{0}$, $p_{0m} = 0$, and $\xi_{0m} = 0$ for all m).

2.2 Cost Function and Supply

I follow Byrne et al. (2022)'s notation to define the cost function. For each product j in market m, researchers observe output q_{jm}, market size, $Q_m = q_{jm}/s_{jm}$, $L \times 1$ vector of input price $\mathbf{w}_m$ and cost C_{jm}. The true cost (in short, cost) C^*_{jm} is assumed to be measured with error, i.e. $C_{jm} = C^*_{jm} + u_{cjm}$, where I assume u_{cjm} is i.i.d. normally distributed. The cost C^*_{jm} is assumed to be a function of output, input prices $\mathbf{w}_m$, observed product characteristics $\mathbf{x}_{jm}$ and a cost shock υ_{jm}. That is,

$$C^*_{jm} = C\left(q_{jm}, \mathbf{w}_m, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right),$$

where $\boldsymbol{\theta}_c$ is a parameter vector. $C(\cdot)$ is assumed to be strictly increasing and continuously differentiable in output and cost shock.

Assuming that there is one firm for each product, firm j's profit function is as follows:

$$\pi_{jm} \quad = \quad p_{jm} \times q_{jm} - C\left(q_{jm}, \mathbf{w}_m, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right).$$

Let MR_{jm} be the marginal revenue of firm j in market m. I follow the literature and assume that firms act as differentiated products Bertrand price competitors. Therefore, the equation below holds from the F.O.C:

$$\begin{aligned} MR_{jm} &= \underbrace{\frac{\partial p_{jm} q_{jm}}{\partial q_{jm}} = p_{jm} + s_{jm}\left[\frac{\partial s_j\left(\mathbf{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \boldsymbol{\theta}_d\right)}{\partial p_{jm}}\right]^{-1}}_{MR_{jm}} \\ &= MC_{jm} = \underbrace{\frac{\partial C\left(q_{jm}, \mathbf{w}_m, \mathbf{x}_{jm}, \upsilon_{jm}; \boldsymbol{\theta}_c\right)}{\partial q_{jm}}}_{MC_{jm}}. \end{aligned} \tag{2}$$

In logit demand specification, it can be expressed as:

$$MR_{jm} = p_{jm} - \frac{1}{\left(1 - s_{jm}\right)\alpha}.$$

I assume the following production function:

$$q = exp\left(x\right)^{-(\alpha_c+\beta_c)/\eta} exp\left(\upsilon\right)^{-(\alpha_c+\beta_c)} L^{\alpha_c} K^{\beta_c},$$

where w represents the labor input cost and r stands for the capital rental rate. The cost and marginal cost functions are given as follows:

$$C^*\left(q, w, r, x, \upsilon\right) = \left[\left(\alpha_c + \beta_c\right)\left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)}\right] exp\left(x\gamma + \upsilon\right) q^{\frac{1}{\alpha_c+\beta_c}} \tag{3}$$

$$MC^*\left(q, w, r, x, \upsilon\right) = \left[\left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)}\right] exp\left(x\gamma + \upsilon\right) q^{\frac{1}{\alpha_c+\beta_c}-1} \tag{4}$$

$$= C^*\left(\right) \times \frac{1}{\left(\alpha_c + \beta_c\right) q}. \tag{5}$$

Throughout this paper, I use the logit demand and cost function derived from Cobb-Douglas production function as the specific functional form example, unless noted otherwise.

2.3 Entry

For now, I assume that in each market, in period 1, there is only one potential entrant with one product. Hence, from now on, whenever I discuss the entry model, I assume that any entrant becomes a monopolist, and thus remove the subscript j. In period 1, the firm observes input prices $\boldsymbol{w}_m$, market size Q_m, observed product characteristics $\mathbf{x}_m$, unobserved product characteristics (i.e, demand shock) ξ_m, cost shock υ_m, entry shock ϵ_{Em}, and the shock of being non-entrant ϵ_{Om}. Based on such information, firm decides whether to enter or not. The decision is based on the anticipated period 2 profit. If the profit is higher than the entry cost plus the entry cost shock minus the shock of staying out, then the firm enters. If otherwise, the firm stays out.

Below, I explain the model in more detail, where I discuss the period 2 profit maximizing behavior of the monopolist first.

First, let the exogenous variables $(\mathbf{w}_m, Q_m, \mathbf{x}_m, \xi_m, \upsilon_m, \epsilon_{Em}, \epsilon_{Om})$ be given and known to firms in both periods 1 and 2. In equilibrium, the profit maximizing monopolist choose p_m so that marginal revenue equals marginal cost:

$$MR_m = \frac{\partial p_m q_m}{\partial q_m} = p_m - \frac{1}{\left(1 - s_m\right)\alpha} = MC_m = \frac{\partial C\left(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c\right)}{\partial q_m} \tag{6}$$

$$q_m = Q_m s_m \tag{7}$$

and the resulting profit is:

$$\pi_m\left(\mathbf{w}_m, Q_m, \mathbf{x}_m, \xi_m, \upsilon_m, \boldsymbol{\theta}\right) = p_m q_m - C\left(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c\right). \tag{8}$$

Next, consider period 1. The firm enters if

$$\pi\left(\mathbf{w}_m, Q_m, \mathbf{x}_m, \xi_m, \upsilon_m, \boldsymbol{\theta}\right) - E_C + \epsilon_{Em} \geq \epsilon_{Om} \tag{9}$$

and stays out if

$$\pi\left(\mathbf{w}_m, Q_m, \mathbf{x}_m, \xi_m, \upsilon_m, \boldsymbol{\theta}\right) - E_C + \epsilon_{Em} < \epsilon_{Om}, \tag{10}$$

where E_C is defined to be the entry cost parameter.

If I assume that both ϵ_{Em} and ϵ_{Om} follow i.i.d. type 1 extreme distribution, then period 1 entry probability can be expressed as follows

$$\begin{aligned} & Pr\Big(\pi\left(\mathbf{w}_m, Q_m, \mathbf{x}_m, \xi_m, \upsilon_m, \boldsymbol{\theta}\right) - E_C + \epsilon_{Em} \geq \epsilon_{Om}\Big) \\ = & \frac{exp\Big(\pi\left(\mathbf{w}_m, Q_m, \mathbf{x}_m, \xi_m, \upsilon_m, \boldsymbol{\theta}\right) - E_C\Big)}{1 + exp\Big(\pi\left(\mathbf{w}_m, Q_m, \mathbf{x}_m, \xi_m, \upsilon_m, \boldsymbol{\theta}\right) - E_C\Big)}. \end{aligned} \tag{11}$$

Next, I describe the orthogonality conditions that I impose for estimation.

2.4 Orthogonality conditions

I assume the orthogonality conditions that the demand and supply shocks, and entry shocks $(\xi_m, \upsilon_m, \epsilon_{Em}, \epsilon_{Om})$ are independent to the input prices, observed product characteristics and the market size, $(\mathbf{w}_m, \mathbf{x}_m, Q_m)$.

$$(\xi_m, \upsilon_m, \epsilon_{Em}, \epsilon_{Om}) \perp\!\!\!\perp (\mathbf{w}_m, \mathbf{x}_m, Q_m) \tag{12}$$

The orthogonality condition is for the population of all firms, which includes both entrants and non-entrants.

2.5 Demand function estimation using instruments.

First, I review the IV based demand estimation by Berry (1994). Berry (1994) assumes that for each market $m = 1, \ldots M$, researchers have data on prices $\mathbf{p}_m$, market shares $\mathbf{s}_m = [s_{0m}, s_{1m},...,s_{J_m m}]'$ and observed product characteristics $\mathbf{X}_m$ for all firms in the market. Then, the market shares are described as follows:

$$s_{jm} = \frac{\exp\left(\mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}\right)}{1 + \sum_{k=1}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)}, j = 1, \ldots, J_m \tag{13}$$

$$s_{0m} = \frac{1}{1 + \sum_{k=1}^{J_m} \exp\left(\mathbf{x}_{km}\boldsymbol{\beta} - p_{km}\alpha + \xi_{km}\right)}, \tag{14}$$

Then, from Equations (13) and (14), I can derive the following equation:

$$\log\left(s_{jm}\right) - \log\left(s_{0m}\right) = \mathbf{x}_{jm}\boldsymbol{\beta} - p_{jm}\alpha + \xi_{jm}. \tag{15}$$

Since the profit maximizing firm tends to have higher prices when the unobserved product quality ξ_{jm} is high, the price and the error are likely to be positively correlated, violating the assumption necessary for the OLS estimated parameters to be unbiased. To deal with this issue, researchers use the instrumental variables (IV) estimation methods.

In the IV method, I assume I have variables $\mathbf{z}_{djm}$ that are orthogonal to the demand shock ξ_{jm}, i.e., $E\left[\xi_{jm}|\mathbf{z}_{djm}\right] = 0$ and are correlated with the price p_{jm}. Then, the orthogonality condition

$$E\left[\xi_{jm}|\mathbf{z}_{djm}\right] = E\left[\left(log\left(s_{jm}\right) - log\left(s_{0m}\right) - \mathbf{x}_{jm}\boldsymbol{\beta} + p_{jm}\alpha\right)|\mathbf{z}_{jm}\right] = 0 \tag{16}$$

identifies the demand parameters $\boldsymbol{\beta}$ and α.

Commonly used examples of the instruments are: input prices $\boldsymbol{w}_m$ in market m, observed product characteristics $\boldsymbol{X}_m$, including those of own and rival firms in the market (BLP instruments), and prices of other markets (Hausman instruments). The logic of the Hausman instruments is that prices of firms in other markets reflect the cost components that are common to firms in all markets, which are assumed to be orthogonal to the demand shocks.

Equations (2) and (6) imply that demand parameters can potentially be identified if there is data on marginal cost[8]. Even without such data, if the cost function is known or can be estimated, I can take its derivative with respect to output to derive the marginal cost. BLP assume that marginal cost is log-linear in output and observed product characteristics, i.e., $MC_{jm} = exp\left(\mathbf{w}_m\boldsymbol{\gamma}_w + \upsilon_{jm}\right)$(see their Equation 3.1). They then use instruments to deal with the endogeneity of output due to cost shocks and of prices due to demand shocks. As long as the parametric specification of the supply side is accurate and there are enough instruments for identification, the demand side orthogonality condition in Equation (16) and the one based on

[8]Genesolve and Mullin (1998) use data on marginal cost to estimate the conduct parameters of the homogeneous goods oligopoly model.

the F.O.C. are sufficient for identifying demand parameters.

2.6 Cost function estimation using instruments

For the sake of simplicity in exposition, consider the following loglinear cost function.

$$lnC^*_{jm} = c_0 + lnq_{jm} \times c_q + lnw_{jm} \times c_w + lnr_{jm} \times c_r + x_{jm}\eta + \upsilon_{jm}$$

As with demand estimation, there are important endogeneity concerns with standard approaches to estimating cost functions. Specifically, output q_{jm} is endogenously determined by profit-maximizing firms as described in Equation (2). That is, all else being equal, less efficient firms tend to produce less. Thus, output is potentially negatively correlated with the cost shock υ_{jm}. Then, the RHS variable can be correlated with the cost shock υ_{jm}, which corresponds to the error term in this equation. Such correlation between the RHS variable and the error term again leads to bias in the OLS estimation. In dealing with this issue, researchers have traditionally focused on selected industries where endogeneity can be ignored, or used instruments for output.

Researchers typically use demand shifters as instruments. Let us denote the vector of cost instruments by $\mathbf{z}_{sjm}$. One can estimate $\boldsymbol{\theta}_c$ by relying on the orthogonality conditions: $E\left[\upsilon_{jm}|\mathbf{z}_{sjm}\right] = \mathbf{0}$. Typical instruments that can be used for both the endogeneity of price in demand function and the endogeneity of output in the cost function are the product characteristics of rival firms in the same market:$\mathbf{X}_m$. In the case of monopoly, I can use the market size Q_m instead.

2.7 Bias due to entry

It is important to note that in conventional demand and cost function estimation, only data on entrant firms are used to form the sample analog of the instrument orthogonality conditions. Then, even if I assume instruments are valid for all firms (i.e., entrant and non-entrant firms), the orthogonality conditions between the unobserved product characteristics and the instruments may fail to approximately hold in sample. The reason is: profitable firms are more likely to enter, producing a correlation between IV and demand/cost shocks among entrant firms. More concretely, let us consider the input price as IV. If the input price is higher in a market than in other markets, given other observed and unobserved variables in the profit function and the shocks ϵ_{Em}, ϵ_{Om} being the same, the firms there will have lower profits than those in other markets, and thus, firms with low unobserved product quality, which would survive in other markets may not do so. Hence, across markets, I should find positive correlation between input price and demand shock for entrants. Next, let us call those firms that have observed product characteristics that are perceived by consumers to be of higher quality. Then, even if it has lower unobserved product characteristics, it would have the same profit, and thus, the same entry probability. Hence, entry behavior creates a negative correlation between the observed product quality and the demand shock for entrants. Furthermore, cost shock in one market is likely to be high if prices in other markets are high, due to the positive correlations in cost shocks across markets. Then, this would result in lower profit than before, and thus, only firms with high unobserved product characteristics would be able to enter, creating the positive correlation between the Hausman instrment (prices of other markets) and the demand shock for entrants.

To illustrate the selection bias resulting from entry and the subsequent bias in demand and cost function estimates generated by Berry (1994) approach, I conduct several Monte-Carlo exercises. To do so, I generated the monopoly equilibrium price, output and profit for each market $m = 1, \dots, M$. Then, using Equation (11), I simulate the entrant firm. The procedure is as follows: Let $\mathcal{E}$ be the set of entrant firms. For each $m = 1, \dots, M$, I draw $\eta_m \in U(0, 1)$. I let $m \in \mathcal{E}$ if

$$\eta_m \leq \frac{exp\left(\pi_m - E_C\right)}{1 + exp\left(\pi_m - E_C\right)},$$

and $m \notin \mathcal{E}$ if otherwise.

In Table 2, I present the sample statistics, first, for the whole sample and then, for entrant firms only. As I can see from Panel A of the table, the sample means and standard deviations of all firms are close to the true values. However, when I look at Panel B, where I present the sample statistics of the entrant firms, both the means and the standard deviations of the demand and cost shocks are quite different from the true values, which indicates the selection bias. In particular, the entrants have, on average, lower wages and rental rates than the nonentrants, and tend to have higher unobserved product characteristics and lower cost than nonentrants. In Panel C, I present the orrelations between the instruments and the demand and cost shocks. In the sample of all firms, the correlations are overall close to zero. However, for the entrants, the correlation between the observed and unobserved product qualities are negative. This is due to the negative correlation between observed and unobserved product qualities for firms with the same profit, and thus, the same entry probability. The positive correlation between the observed product characteristics and the cost shock for the entrants occurs for the same reason. The reason for the correlations between the input prices and the unobserved product characteristics being positive is the same as well. These results suggest that if I restrict the sample to entrants, the IV orthogonalities fail to hold.

I then use the above artificially generated sample and derive the OLS estimates and the IV estimates using valid instruments, i.e. instruments that satisfy the instrument orthognality conditions for the population of all firms, including non-entrants. In Table 3, I report the parameter estimates. In Column (1), I present the OLS results and in Column (2), the IV results when I use all firms. We can see that on average, the OLS estimated price parameter α is downwardly biased (i.e. the price coefficient $-\alpha$ is upwardly biased). This is likely due to the price endogeneity. However, when I used input prices and observed product characteristics as instruments, the IV estimates are close to the true values, which demonstrates the validity of the instruments. I then report the results in Column (3), where I only used the entrants as sample. Then, both the price parameter and the parameter of the observed product characteristics are downwardly biased. The source of bias for the price parameter is the positive correlation between the price and the demand shock, and the positive correlation between the input price and the observed product characteristics for entrant firms. Since the input prices and the output prices are positively correlated, these correlation positively bias the price coefficient(i.e, negative bias on the price parameter). In contrast, the observed product characteristics and the demand shock are negatively correlated, resulting in the downward bias for the coefficient on the observed product characteristics. Next, I report the same for the Cobb-Douglas cost function. The estimating equation is:

$$lnC^* = Const + \frac{\alpha_c}{\alpha_c + \beta_c} lnw + \frac{\beta_c}{\alpha_c + \beta_c} lnr + \frac{1}{\alpha_c + \beta_c} lnq + \eta x + \upsilon$$

In Table 3, I report the parameter estimate α_c and β_c. I can see that the OLS estimates in Column (1) are biased even if I use the simulated data for all firms, suggesting bias due to the endogeneity of output. In contrast, the IV estimates in Column (2) are close to the true values when I use simulated data for all firms, which again indicates the validity of the instruments. However, the IV estimates have large bias if I use only data on entrant firms, as shown in Column (3). The negative correlation between the output and the cost shock, shown in Panel B in Table 2, results in the downward bias of the output coefficients, and thus, upward bias of the estimate of $\alpha_c + \beta_c$, which results in upward bias of the parameter estimates of α_c and β_c.

2.8 Estimation of Static Games

Ciliberto and Tamer (2009) considered the estimation of static entry games with multiple equilibria. Their contribution is that they derived the bounds of the entry probabilities without imposing an arbitrary equilibrium selection rule. The basic idea of their methodology is, given the parameters, first draw the entry shock for all entrants and non-entrants, compute all Nash equilibria of the entry game. Then, for each combination of entrant firms, they derive the lower bound of the entry probability. They do so by only choosing the nonentry equilibrium in multiple equilibria situations. The upper bound of the entry probability can be computed similarly. They then use these lower and upper bounds to estimate the bounds of the parameters of the entry model.

Ciliberto et al. (2021) also use similar ideas. They add to the framework of Ciliberto and Tamer (2009) the 2nd stage where the entrants play the differentiated products oligopoly pricing game. They also modify the first stage and assume firms anticipate the equilibrium profit of the 2nd stage pricing game, and decide whether to enter or not. Since the uniqueness of price competition under the logit demand model is proven by Mizuno (2003), they ensure that the issue of multiple equilibria only arises in the first-stage entry game. In the estimation of the bounds of the parameters, they follow Ciliberto and Tamer (2009).

I first present a monopoly version of the entry model by Ciliberto et al. (2021). It consistes of the following three equations, and one moment condition.

1 Logit demand:

$$lns_m - lns_{0m} = -p_m\alpha + \mathbf{x}_m\boldsymbol{\beta} + \xi_m \tag{17}$$

2 F.O.C. of profit maximization:

$$p_m - \frac{1}{(1-s_m)\,\alpha} = MC = exp\left(\mathbf{w}_m\boldsymbol{\varphi} + \mathbf{x}_m\boldsymbol{\gamma} + \upsilon_m\right) \tag{18}$$

3 Entry

$$I_{em} = \begin{cases} 1 & if\ \Big(p_m - exp\left(\mathbf{w}_m\boldsymbol{\varphi} + \mathbf{x}_m\boldsymbol{\gamma} + \upsilon_m\right)\Big)Q_m s_m - exp\left(\mathbf{z}_m\boldsymbol{\gamma}\right) + \epsilon_{Em} > \epsilon_{Om} \\ 0 & if\ otherwise \end{cases}. \tag{19}$$

4 Orthogonality condition.

$$(\xi_m, \upsilon_m, \epsilon_{Em}, \epsilon_{Om}) \perp\!\!\!\perp (\mathbf{w}_m, \mathbf{x}_m, Q_m, \mathbf{z}_m)\,. \tag{20}$$

Equation (17), is the logit demand equation. Equation (18) states that firms equate marginal revenue to marginal cost. Its LHS is the marginal revenue of the logit market share demand, and the RHS is the marginal cost function, which is specified as the log linear function of the input price $\mathbf{w}_m$ and the cost shock υ_m.

Equation (19) describes the entry decision of the firm. The first term of the RHS is the profit, where $Q_m s_m$ equals q_m, the output, and the 2nd term is the entry cost. The entry cost is a log linear function of the observed variables $\mathbf{z}_m$, which affects entry but does not enter in the profit function, and thus, works as instruments for the endogenous entry choice of the firm. ϵ_{Em}, ϵ_{Om} are the idiosyncratic shocks to the benefit of entering and staying out, respectively. $I_{em} = 1$ if the firm is an entrant, i.e., if the benefit of entry is higher than the benefit of staying out, and $I_{em} = 0$ if otherwise. While the exogenous variables $\mathbf{w}_m$, Q_m, $\mathbf{z}_m$ and $\mathbf{x}_m$ are observable to the researcher for all firms, including both entrants and non-entrants, the endogenous variables s_m, s_{0m}, p_m are only observable for entrants, i.e., firms with $I_{em} = 1$. Furthermore, ξ_m, υ_m and $\epsilon_{Em} - \epsilon_{Om}$ are unobservable for all firms. They are assumed to be independent to $\mathbf{w}_m$, Q_m, $\mathbf{z}_m$ and $\mathbf{x}_m$, and Q_m.

In the literature, the identification of the above model relies on the orthogonality condition in Equation (20). To use the moment condition effectively, the literature, such as Ciliberto et al. (2021) and Aguirregabiria et al. (2024), assume that all the exogenous variables on the RHS are observable.[9] Then, given the exogenous variables and the orthogonality condition, given the parameters of the unobserved variables on the LHS of the orthogonality condition, Ciliberto et al. (2021) draw the unobserved variables, and compute the price, market share and other variables for the entrant firms, and the entry probability. Then, they compare the simulated moments of those variables to the ones in the data, and choose the parameters that makes the two sets of moments the closest.

In the following paragraph, I replicate their methodology on the above entry model of monopolists. More concretely, given the parameters of the model, I can recover ξ_m, υ_m of the entrant firms from Equations (17) and (18), and thus their empirical joint distribution of the entrant

[9] Note that they do not use market size Q_m.

firms as well. Let us denote the empirical distribution of (ξ_m, υ_m) of the entrant firms by

$$F^d \left(\xi_m \left(s_m, p_m, \mathbf{x}_m\right), \upsilon_m \left(s_m, p_m, \mathbf{w}_m\right) | I_{em} = 1, \mathbf{w}_m, Q_m, \mathbf{z}_m, \mathbf{x}_m; \alpha, \boldsymbol{\beta}, \boldsymbol{\varphi}\right). \tag{21}$$

Next, given the parameters of the model and the exogenous variables $(\mathbf{w}_m, Q_m, \mathbf{z}_m \mathbf{x}_m)$, generate the equivalent distribution by simulation as follows:

Step 1 First, draw $(\xi_m, \upsilon_m, \epsilon_{Em}, \epsilon_{Om})$, using the orthogonality conditions in Equation (20), where $(\mathbf{w}_m, \mathbf{x}_m, Q_m, \mathbf{z}_m)$ are observed in the data.

Step 2 Solve the profit maximization problem to derive s_m, s_{0m}, p_m for all firms given $\mathbf{w}_m$, Q_m, $\mathbf{z}_m$ and $\mathbf{x}_m$, and (ξ_m, υ_m) drawn in Step 1. Then, derive the profit.

Step 3 Given the profit π_m, simulate $(\epsilon_{Em}, \epsilon_{Om})$ and derive I_{em}.

After repeated calculations, I can derive the simulated distribution of (ξ_m, υ_m) entrants.

$$F^s \left(\xi_m \left(s_m, p_m, \mathbf{x}_m\right), \upsilon_m \left(s_m, p_m, \mathbf{w}_m\right) | I_{em} = 1, \mathbf{w}_m, Q_m, \mathbf{z}_m, \mathbf{x}_m; \alpha, \boldsymbol{\beta}, \boldsymbol{\varphi}, \boldsymbol{\gamma}, f_{(\xi,\upsilon)}, f_{(\epsilon_E,\epsilon_O)}\right). \tag{22}$$

Then, parameters can be estimated by choosing the ones that makes the the two distributions defined in (21) and (22) as close as possible.

Aguirregabiria et al. (2024) uses the moment condition in Equation (20) to construct the entry probability given the RHS variables, and then, after estimating the entry probability from the sample frequency of entry conditional on the RHS variables, and then, constructs the population moments which equals the sample moment weighted by the inverse of the entry probability.

Both identification strategies require observability of $(\mathbf{w}_m, \mathbf{x}_m, Q_m, \mathbf{z}_m)$. In particular, it requires that the product characteristics $\mathbf{x}_m$ is observable for both entrants and non-entrants. In many industries, we know very little about the potential entrants, including their product characteristics $\mathbf{x}_m$. The exception is the airline industry, in which a market is a specific route from an airport to another, and the potential entrants are existing airliners, whose product characteristics $\mathbf{x}_m$ are known. Since they already operate in other routes, one could consider them as entrants in the airline industry, thus not as potential entrants.

Now, consider the case where the product characteristics $\mathbf{x}_m$ are observable for the entrants, but unobservable for the non-entrants. Then, even though I can derive the empirical distribution in Equation (21), I cannot do so for the distribution in Equation (22). The reason is that in order to derive the profit in step 2 for each each firm, including the nonentrant, I need to draw the missing $\mathbf{x}_m$ for the non-entrants in the data given the observable exogenous variables $(\mathbf{z}_m, \mathbf{w}_m, Q_m)$. However, I do not have any restrictions from the model that I can use to draw $\mathbf{x}_m$ for the non-entrants, other than that they are orthogonal to the unobserved $(\xi_m, \upsilon_m, \epsilon_{Em}, \epsilon_{Om})$.

Without observing the product characteristics $\mathbf{x}_m$ for all firms including nonentrants, I cannot conduct Step 1 of the simulation procedure, and thus, cannot derive the joint distribution of (ξ_m, υ_m) for the entrants generated from the model, and compare the joint distribution derived from the data on entrants, and choose the parameters that make those two distributions as close as possible.

For similar reasons, without observing $\mathbf{x}_m$ for all firms that includes non-entrants, we cannot estimate the entry probability as a function of the observable exogenous variables ($\mathbf{x}_m$,$\mathbf{z}_m$,$\mathbf{w}_m$,Q_m), as is done in Aguirregabiria et al. (2024) and others in the literature as the first step, which then is used for the sample selection bias correction, or to re-weight the moments for unbiased estimation using only entrants.

Next, I discuss two different ways for the identification and estimation in such a situation. First, since $\mathbf{x}_m$ is unobservable for the non-entrant, I propose an identification strategy, i.e., orthogonality condition that conditions on the product characteristics $\mathbf{x}_m$ of the entrant firm. I show that identification can be established, but since it rests on the joint distribution of $(\mathbf{s}_m, \mathbf{p}_m)$ being the same conditional on a specific variation of $\mathbf{w}_m$ given $\mathbf{x}_m$, it is likely to be impractical for the estimation of the structural parameters unless the sample size is extremely large. I show in my Monte-Carlo exercises that significant biases of the parameter estimates remain even in large sample size.

In the other identification strategy, I apply the recent results by Byrne et al. (2022) and Imai et al. (2024) that use cost data to estimate the input price coefficients in cost function and, in particular, the price coefficients in demand function and the output coefficients in cost function that are subject to endogeneity problems. Since the estimation does not require any instruments, I only need data on entrants to obtain consistent estimates of those key demand and cost function parameters. As I have already shown in Column (3) of Table 3, conventional IV estimates do not provide consistent results, due to the endogeneity of entry. I present the results in Table 4 where I use the same data as those used in Column (3) of Table 3, but apply the instrument-free procedure of Imai et al. (2024) using cost data. I can see that it delivers consistent parameter estimates.

In the next section, I review the recent instrument-free approaches in demand and cost function estimation, proposed by Byrne et al. (2022) and Imai et al. (2024)

3 Instrument-free identification of demand and cost functions using cost data

This section outlines the method for estimating the demand and cost function parameters $(\alpha, \alpha_c, \beta_c)$ of the logit monopoly entry model, by applying the estimation strategy discussed in Byrne et al. (2022) and Imai et al. (2024), who use cost data instead of instruments for their identification. I do so before explaining the identification without the cost data because it also uses some of their results. Their approach offers two main advantages. First, their model inherently cancels out the cost shock in the estimation equation, eliminating the need for instrumental variables. Second, the parameters can be estimated using data solely on entrants. This method addresses selection bias due to entry, as the selection stems from unobserved demand and cost shocks, which their model already accounts for. Consequently, researchers only need to fit the modified parametric cost function to the observed costs. Extensions to oligopoly model and random coefficient model discussed in Berry et al. (1995) are described in section 5.

To begin with, let $\mathcal{E}$ be the set of markets with monopolist entrant. Then, I assume only variables of firms in $\mathcal{E}$ are observable. Then, the cost function can be expressed as follows:

$$C_m = C_m^* + u_{cm} = C^* \left(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_{c0}\right) + u_{cm},\ m \in \mathcal{E}$$

where C_m^* is the true cost and u_{cm} is the measurement error or the idiosyncratic component of cost. I assume the error term u_{cm} to be independent to the output q_m, input price vector $\mathbf{w}_m$, vector of observed product characteristics $\mathbf{x}_m$, and the demand variables, which are: price p_m and market share s_m. Then, assume that the cost function is specified as follows:

$$C^* \left(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c\right) = \widetilde{C} \left(q_m, \mathbf{w}_m, \mathbf{x}_m; \boldsymbol{\theta}_c\right) \varphi \left(\mathbf{x}_m, \upsilon_m\right).$$

Then, taking the marginal cost, I derive

$$MC^* \left(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c\right) = \widetilde{MC} \left(q_m, \mathbf{w}_m, \mathbf{x}_m; \boldsymbol{\theta}_c\right) \varphi \left(\mathbf{x}_m, \upsilon_m\right),$$

where $\widetilde{MC}\left(\right)$ is the derivative of the function $\widetilde{C}\left(\right)$ with respect to the output q. Then,

$$\frac{C^* \left(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c\right)}{MC^* \left(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c\right)} = \frac{\widetilde{C} \left(q_m, \mathbf{w}_m, \mathbf{x}_m; \boldsymbol{\theta}_c\right)}{\widetilde{MC} \left(q_m, \mathbf{w}_m, \mathbf{x}_m; \boldsymbol{\theta}_c\right)},$$

which does not depend on the supply shock υ_m.

If I assume Cobb-Douglas production function as in Equation (3), then from Equations (3) and (5), I derive

$$\frac{C^*(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c)}{MC^*(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c)} = (\alpha_c + \beta_c)\, q_m,\ m \in \mathcal{E}. \tag{23}$$

Note that the RHS does not contain the unobservable cost shock υ_m. Furthermore, from the F.O.C.,

$$MR(p_m, s_m, \mathbf{x}_m; \boldsymbol{\theta}_d) = MC^*(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c), m \in \mathcal{E}. \tag{24}$$

Under the logit demand assumption,

$$MR(p_m, s_m, \mathbf{x}_m; \boldsymbol{\theta}_d) = p_m - \frac{1}{(1 - s_m)\,\alpha}. \tag{25}$$

Using Equations (24), (25) to substitute $MR()$ for $MC()$ in Equation (23), I obtain

$$\begin{aligned} C^*(q_m, \mathbf{w}_m, \mathbf{x}_m, \upsilon_m; \boldsymbol{\theta}_c) &= \frac{\widetilde{C}(q_m, \mathbf{w}_m, \mathbf{x}_m; \boldsymbol{\theta}_c)}{\widetilde{MC}(q_m, \mathbf{w}_m, \mathbf{x}_m; \boldsymbol{\theta}_c)} MR(p_m, s_m, \mathbf{x}_m; \boldsymbol{\theta}_d) & (26) \\ &= (\alpha_c + \beta_c)\, q_m \left(p_m - \frac{1}{(1 - s_m)\,\alpha}\right). & (27) \end{aligned}$$

This is how Imai et al. (2024) expressed the cost function as a function that does not have the unobservable cost shock υ_m, which was the source of the endogeneity bias.

Since the observed cost is the true cost function plus the measurement error, it can be specified as follows:

$$C_m = p_m q_m (\alpha_c + \beta_c) - \frac{q_m}{1 - s_m} \frac{\alpha_c + \beta_c}{\alpha} + u_{cm}.\ m \in \mathcal{E} \tag{28}$$

I also use the Shephard's lemma, which states:

$$\frac{\partial ln C^*(q_m, \mathbf{w}_m, \mathbf{x}_m; \boldsymbol{\theta}_c)}{\partial ln w_{lm}} = \frac{w_{lm} L_{lm}}{C^*_m},\ l = 1, \ldots, L \tag{29}$$

In the Cobb-Douglas cost function example, $\mathbf{w}_m = (w_m, r_m)$, where w_m is the input price for labor input, which I denote as L_m, and r_m is the input price for capital input,which I denote as K_m. During the estimation, they only use the first $L-1$ of the above equation because the sum of L cost shares $w_{lm} L_{lm} / C^*_m$ adds up to one. Substituting C^*_m in Equation (27) into the Equation (29), I derive for the labor input,

$$w_m L_m = \alpha_c q_m \left(p_m - \frac{1}{(1 - s_m)\,\alpha}\right). \tag{30}$$

I follow Imai et al. (2024) and assume that input cost is measured with error. Hence, the labor input cost C_{Lm} is specified as

$$C_{Lm} = p_m q_m \alpha_c - \frac{q_m}{1 - s_m} \frac{\alpha_c}{\alpha} + u_{Lm}, \ m \in \mathcal{E} \tag{31}$$

where u_{Lm} is the measurement error of the labor input cost.

In sum, the RHSs of the Equations (28) or (31) do not contain either cost shock or demand shock. Their error terms are measurement errors that are assumed to be independent to all observed and unobserved exogenous variables. Hence, the measurement errors are independent to the endogenous RHS variables, even if I restrict the sample to entrant firms. Furthermore, price enters in the first term of the RHS but not in the second term, and market share $s_{jm} = q_{jm}/Q_m$ enters in the second term but not in the first. Those restrictions ensure that the two terms are not collinear. Hence, I can obtain unbiased estimates of $\alpha_c + \beta_c$ and $(\alpha_c + \beta_c)/\alpha$ from Equation (28) and unbiased estimates of α_c and α_c/α from Equation (31), thus, consistent estimates of α_c, β_c and α.

4 Identification and estimation of entry cost and other parameters under monopoly.

4.1 Introduction to identification

In this section, I first discuss identification of the model parameters, which includes the price parameter α, the Cobb-Douglas production function parameters α_c and β_c, the entry cost parameter E_C, the coefficients of the observed product characteristics $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, and the joint distribution of $\mathbf{x}$, ξ, and v in the monopoly entry model I described in Subsection 2.3. In each market m, I assume there is only one potential entrant that decides whether to enter in market m as a monopolist or not. To begin with, I first derive a few identities:

$$
\begin{aligned}
\ln s_m - \ln(1-s_m) &= -p_m\alpha + \mathbf{x}_m\boldsymbol{\beta} + \xi_m && (32)\\
\ln\left[p_m - \frac{1}{(1-s_m)\,\alpha}\right] &= \frac{\alpha_c}{\alpha_c+\beta_c}\left(\ln w_m - \ln\alpha_c\right) + \frac{\beta_c}{\alpha_c+\beta_c}\left(\ln r_m - \ln\beta_c\right)\\
&\quad + \left(\frac{1}{\alpha_c+\beta_c} - 1\right)\ln\left(Q_m s_m\right) + \mathbf{x}_m\boldsymbol{\gamma} + \upsilon_m\\
&= \frac{\alpha_c}{\alpha_c+\beta_c}\left(\ln w_m - \ln\alpha_c\right) + \frac{\beta_c}{\alpha_c+\beta_c}\left(\ln r_m - \ln\beta_c\right)\\
&\quad + \left(\frac{1}{\alpha_c+\beta_c} - 1\right)\ln\left(s_m\right) + \left(\frac{1}{\alpha_c+\beta_c} - 1\right)\ln\left(Q_m\right) + \mathbf{x}_m\boldsymbol{\gamma} + \upsilon_m && (33)\\
\pi_m = p_m q_m - C_m^* &= q_m\left[\left(1-\alpha_c-\beta_c\right)p_m + \frac{\alpha_c+\beta_c}{\alpha}\frac{1}{(1-s_m)}\right]\\
&= s_m Q_m\left[\left(1-\alpha_c-\beta_c\right)p_m + \frac{\alpha_c+\beta_c}{\alpha}\frac{1}{(1-s_m)}\right] && (34)\\
P_E\left(\pi_m\right) &= \Pr\left(\pi_m - E + \epsilon_E \geq 0\right) && (35)
\end{aligned}
$$

Next, I first consider the case where I do not have cost data, i.e., I only have data on $\mathbf{x}_m$, p_m, s_m, and $q_m = Q_m s_m$ for the entrant, and assume the demand function to be logit, as specified in Equations (13) and (14) and the production cost function to be Cobb Douglas, as specified in Equation (3).

4.1.1 Identification without cost data

Below, I propose to construct moments that only requires product characteristics for entrant firms.

Proof of Identification: Please refer to Appendix 9A.

Even though I showed that parameters are identified without the cost data, later Monte-Carlo studies I provide demonstrate large biases of the estimated parameter $R = \alpha_c / \left(\alpha_c + \beta_c\right)$. I claim that this is due to the source of the variation that identifies the parameter R being *conditional* on $\mathbf{x}_m$. That is, R_0 is identified from (p_m, s_m) generated from any (w_m, r_m) satisfying $Rw_m + (1-R)\,r_m = A$ conditional on $\mathbf{x}_m = \mathbf{x}$ having the same distribution. For such an identification argument to work in the actual data, I would need large sample of data $(w_m, r_m, \mathbf{x}_m)$ being close to $(w, r, \mathbf{x})$. Monte-Carlo experiments demonstrate that large sample and large variation of (w_m, r_m) are necessary for identification.

Another issue is the unobservability of the number of potential entrants. Since potential entrants are not observable, their number in a market cannot be observed either. Then, the entry probability needs to be modified to include the number of potential entrants as the additional variable. Then, if the number of potential entrant is correlated with the input prices, $R_0 \equiv \alpha_{c0}/\left(\alpha_{c0}+\beta_{c0}\right)$ cannot be identified.

That is why I also discuss in the next subsection the identification of the model in markets where I have cost data. I will see that in that case, identification requires far smaller sample size, and even if we have variation in the number of potential entrants, which is unobservable and correlated with the input price, as long as we have the cost data, we can identify the price coefficient α and the production function parameters α_c and β_c.

4.1.2 Identification using cost data.

Proof of Identification: Please refer to Appendix 9B.

Remarks: 1) As the above proof shows, for identification I only need data on entrant firms. 2) Obtaining the true cost C_m^* of the entrants in the first stage allows us to derive the profit π_m for the entrants. Then, given π_m, I can condition on the entry probability, and thus, control for the sample selection bias.

4.2 Estimation issues

An estimation procedure that reflect the above identification analysis is to estimate based on the sample orthogonality conditions constructed from

$$\frac{f\left(\mathbf{x},\delta,\xi,1|\mathbf{w},Q\right)}{Pr\left(\pi-E_C\right)}=f_{\mathbf{x}}\left(\mathbf{x}|\mathbf{w},Q\right)f_{(\xi,\upsilon)}\left(\delta-\mathbf{x}\boldsymbol{\beta},\eta-\mathbf{x}\boldsymbol{\gamma}\right).$$

However, in Monte-Carlo studies, I find that simple weighting by the inverse of the entry probability results in instability of the parameter estimates due to high weight put on the sample with low entry probability. To deal with this issue, I needed to trim away those samples, which in some cases resulted in large bias of the parameter estimates. I also anticipate that such bias problems to become more severe when I apply the approaches based on the inverse probability weights to the estimation of the oligopoly model. Hence, I instead use the simulation based method that randomly generates the missing data on non-entrant firms, i.e. MCMC estimation with data augmentation.

4.3 MCMC Estimation with Data Augmentation

I employ MCMC with data augmentation. I use uninformative prior so that the posterior is the likelihood.

I first construct the log likelihood assuming that I observe data on non-entrants as well as entrants. Then, the likelihood is based on the distribution of unobservable exogenous variables $(u_{cm}, u_{lm}, \xi_m, \upsilon_m, \epsilon_{Em}, \epsilon_{Om})$. I assume that $\mathbf{u}_m \equiv (u_{cm}, u_{lm})' \sim N(\mathbf{0}, \mathbf{\Sigma_u})$, $\xi_m \sim N(\mu_\xi, \sigma_\xi)$, $\upsilon_m \sim N(\mu_\upsilon, \sigma_\upsilon)$, and $(\epsilon_{Em}, \epsilon_{Om})$ are i.i.d. type 1 extreme value distributed. I denote

$$\begin{aligned} \Sigma_{\mathbf{u}11} &= \sigma_c^2 \\ \Sigma_{\mathbf{u}22} &= \sigma_l^2 \\ \Sigma_{\mathbf{u}12} = \Sigma_{\mathbf{u}21} &= \rho\sigma_c\sigma_l \end{aligned}$$

Let l_{um} be the part of the log likelihood increment that is based on the measurement error of cost. Then,

$$l_{um} = -ln\pi - \frac{1}{2}ln\left|\mathbf{\Sigma_u}\right| - \frac{1}{2}\mathbf{u}_m'\mathbf{\Sigma_u}^{-1}\mathbf{u}_m, . \tag{36}$$

where

$$C_m^* = (\alpha_c + \beta_c)\, q_m \left(p_m - \frac{1}{(1-s_m)\,\alpha}\right) \tag{37}$$

$$u_{1m} = u_{cm} = C_m - C_m^* \tag{38}$$

$$u_{2m} = u_{Lm} = C_{Lm} - \frac{\alpha_c}{\alpha_c + \beta_c}C_m^* \tag{39}$$

Next, let l_{hm} be the part of the log likelihood increment that is based on the utility and cost shocks for all firms including non-entrants. In the monopoly model explained above, (δ_m, η_m) is a function of price and market share (p_m, s_m) given the observable exogenous variables (w_m, r_m, Q_m). I denote it as

$$(\delta_m, \eta_m) = h\left((s_m, p_m)\,|w_m, r_m, Q_m\right).$$

More concretely, the firm specific exogenous component of utility is

$$\delta_m = h_1\left((s_m, p_m)\,|w_m, r_m, Q_m\right) = ln\left(s_m\right) - ln\left(1 - s_m\right) + p_m\alpha \tag{40}$$

and the unobservable component of the cost shock is, from the F.O.C.

$$\begin{aligned}\eta_m &= h_2\left(s_m, p_m | w_m, r_m, Q_m\right) = ln\left(p_m - \frac{1}{\left(1-s_m\right)\alpha}\right) \\ &-\Big[\frac{\alpha_c}{\alpha_c+\beta_c} ln\left(\frac{w_m}{\alpha_c}\right) + \frac{\beta_c}{\alpha_c+\beta_c} ln\left(\frac{r_m}{\beta_c}\right) \\ &+\left(\frac{1}{\alpha_c+\beta_c} - 1\right)\left(lnQ_m + lns_m\right)\Big]. \end{aligned} \tag{41}$$

Therefore, the likelihood of $\mathbf{v}_m = (s_m, p_m)$ conditional on $(\mathbf{x}_m, w_m, r_m)$ is

$$f\left(\left(s_m, p_m\right)|\mathbf{x}_m, w_m, r_m, Q_m; \boldsymbol{\theta}\right) = f_{(\xi,\upsilon)}\left(\left[h\left[\left(s_m, p_m\right)|w_m, r_m, Q_m\right] - \left(\mathbf{x}_m\boldsymbol{\beta}, \mathbf{x}_m\boldsymbol{\gamma}\right)\right]\right) J_{((\xi_m,\upsilon_m)\to(s_m,p_m))}$$

where

$$J_{((\xi,\upsilon)\to(s_m,p_m))} = \left\|\frac{\partial\left(\xi_m, \upsilon_m\right)}{\partial\left(s_m, p_m\right)}\right\|$$

is the Jacobian. The use of Jacobian in likelihood evaluation for MCMC can be seen in Jiang et al. (2009). The Jacobian can be derived as follows:

$$\begin{aligned} d\xi &= \left[\frac{1}{s} + \frac{1}{1-s}\right] ds + \alpha dp \\ d\upsilon &= \left[-\frac{1}{MR\left(1-s\right)^2\alpha} - \left(\frac{1}{\alpha_c+\beta_c} - 1\right)\frac{1}{s}\right] ds + \frac{1}{MR} dp \\ MR &= p - \frac{1}{\left(1-s\right)\alpha}. \end{aligned}$$

Hence,

$$J_{((\xi,\upsilon)\to(s,p))} = \left\|\begin{array}{cc} \frac{1}{s} + \frac{1}{1-s} & \alpha \\ -\frac{1}{MR(1-s)^2\alpha} - \left(\frac{1}{\alpha_c+\beta_c} - 1\right)\frac{1}{s} & \frac{1}{MR} \end{array}\right\| \tag{42}$$

Thus, the second component of the log likelihood increment for firm m, which is based on the likelihood of (s_m, p_m) conditional on the observable exogenous variables, is as follows:

$$\begin{aligned} l_{hm} &\equiv lnf\left(\left(s_m, p_m\right)|\mathbf{x}_m, w_m, r_m, Q_m; \boldsymbol{\theta}\right) \\ &= lnf_{(\xi,\upsilon)}\left(h\left(\left(s_m, p_m\right)|w_m, r_m, Q_m\right) - \left(\mathbf{x}_m\boldsymbol{\beta}, \mathbf{x}_m\boldsymbol{\gamma}\right)\right) + ln\left(J_{((\xi_m,\upsilon_m)\to(s_m,p_m))}\right). \end{aligned} \tag{43}$$

The third component of the log likelihood, which I denote as l_{em} is based on the entry probability. Since I assume that the entry shocks $(\epsilon_{Em}, \epsilon_{Om})$ are i.i.d. type I extreme distribution function, the entry probability can be expressed as

$$P_E\left(p_m q_m - C_m^* - E_C\right) = Pr\left(p_m q_m - C_m^* - E_C + \epsilon_{Em} \geq \epsilon_{Om}\right) = \frac{exp\left(p_m q_m - C_m^* - E_C\right)}{1 + exp\left(p_m q_m - C_m^* - E_{C0}\right)}.$$

Hence,

$$l_{em} = ln\left(\frac{exp\left(p_m q_m - C_m^* - E_C\right)}{1 + exp\left(p_m q_m - C_m^* - E_C\right)}\right) I\left(m \in \mathcal{E}\right) + ln\left(\frac{1}{1 + exp\left(p_m q_m - C_m^* - E_C\right)}\right) I\left(m \notin \mathcal{E}\right).$$

Finally, the fourth component of the log likelihood is the distribution of $\mathbf{x}_m$ conditional on the exogenous variables (w_m, r_m, Q_m) and the parameter vector of the distribution, $\boldsymbol{\theta}_{\mathbf{x}}$:

$$l_{\mathbf{x}m} = lnf_{\mathbf{x}}\left(\mathbf{x}_m | w_m, r_m, Q_m, \boldsymbol{\theta}_{\mathbf{x}}\right).$$

Adding all components of the log likelihood, I derive the following log likelihood increments for firm m. Let $\mathbf{u}_m$ be defined as in Equations (38), (39) and C_m^* as in Equation (37). Then,

$$\begin{aligned}
& l_m\left(s_m, p_m, w_m, r_m, Q_m, C_m, C_{Lm}, \mathcal{E}\right) \\
\equiv\ & l_{um} + l_{hm} + l_{em} + l_{\mathbf{x}m} \\
=\ & \left[-\frac{K}{2} ln\pi - \frac{1}{2} ln\left|\boldsymbol{\Sigma}_{\mathbf{u}}\right| - \frac{1}{2}\mathbf{u}_m' \boldsymbol{\Sigma}_{\mathbf{u}}^{-1} \mathbf{u}_m\right] I\left(m \in \mathcal{E}\right) \\
& + lnf_{(\xi,\upsilon)}\left(h\left(\left(s_m, p_m\right) | w_m, r_m, Q_m\right) - \left(\mathbf{x}_m \boldsymbol{\beta}, \mathbf{x}_m \boldsymbol{\gamma}\right)\right) + ln\left(J_{\left(\left(\xi_m, \upsilon_m\right) \to \left(s_m, p_m\right)\right)}\right) \\
& + ln\left(\frac{p_m q_m - C_m^* - E_C}{1 + exp\left(p_m q_m - C_m^* - E_C\right)}\right) I\left(m \in \mathcal{E}\right) + ln\left(\frac{1}{1 + exp\left(p_m q_m - C_m^* - E_C\right)}\right) I\left(m \notin \mathcal{E}\right). \\
& + lnf_{\mathbf{x}}\left(\mathbf{x}_m | w_m, r_m, Q_m, \boldsymbol{\theta}_{\mathbf{x}}\right)
\end{aligned} \tag{44}$$

Note that for l_{um}, l_{hm}, l_{em}, and $l_{\mathbf{x}m}$, the likelihood increments includes non-entrant firms, whose market level data, i.e., input prices (w_m, r_m, Q_m) are observable but firm-level data $(\mathbf{x}_m, s_m, p_m, C_m, C_{Lm})$ are not available. I use data augmentation techniques to simulate the missing exogenous variable $\mathbf{x}_m$ and then, (s_m, p_m) for the non-entrant firms. Given the parameters, observables (w_m, r_m, Q_m) and the augmented $(\mathbf{x}_m, s_m, p_m)$, I can derive the unobservable shocks (ξ_m, υ_m), and then, the true cost C_m^*. In principle, I can also draw the measurement error vector (u_{1m}, u_{2m}) for the potential entrants as well. Note, however, that the measurement errors are independent to any exogenous and endogenous variables, and thus, independent to the entry decision. Hence, drawing measurement errors does not help in the estimation of any parameters of interest. Therefore, I do not augment the missing measurement errors of the non-entrants. In the following, I discuss the data augmentation in more detail.

4.3.1 Data Augmentation

Given (w_m, r_m, Q_m), we explain the procedure at iteration t. At the beginning of iteration t (or at the end of iteration $t-1$), we have $\left(\mathbf{x}_m^{(t-1)}, s_m^{(t-1)}, p_m^{(t-1)}\right)$, and the parameter vector $\boldsymbol{\theta}^{(t-1)}$ which includes $\left(\boldsymbol{\beta}^{(t-1)}, \boldsymbol{\gamma}^{(t-1)}, \boldsymbol{\theta}_{\mathbf{x}}^{(t-1)}\right)$ from iteration $t-1$. I then generate $\left(\mathbf{x}_m^{(t)}, s_m^{(t)}, p_m^{(t)}\right)$ and $\boldsymbol{\theta}^{(t)}$. Since (s_m, p_m) and product characteristics $\mathbf{x}_m$ of entrants are observed, at iteration t, I set

$$\left(\mathbf{x}_m^{(t)}, s_m^{(t)}, p_m^{(t)}\right) = (\mathbf{x}_m, s_m, p_m)\,,\ m \in \mathcal{E}.$$

I then sample the unobserved $\left(\mathbf{x}_m^{(t)}, s_m^{(t)}, p_m^{(t)}\right)$ for nonentrants. I do so by repeating the below procedure for each nonentrant firm $m \notin \mathcal{E}$.

1. Sampling $\mathbf{x}_m$ for $m \notin \mathcal{E}$: I first augment the missing $\mathbf{x}_m^{(t)}$. To do so, I use the Random walk Metropolis-Hastings algorithm, where, from $k=1$ to K, I successively draw the proposal value for the kth element of the vector $\mathbf{x}_m^{(t)}$ given others, to be the same. Now, denote the components of the likelihood increment that includes $\mathbf{x}$ to be

$$\begin{aligned} & l_{\mathbf{x}}\left(s_m^{(t-1)}, p_m^{(t-1)}, w_m, r_m, Q_m, \mathbf{x}; \boldsymbol{\theta}^{(t-1)}\right) \\ \equiv\ & lnf_{(\xi,\upsilon)}\left(h\left(\left(s_m^{(t-1)}, p_m^{(t-1)}\right)|w_m, r_m, Q_m\right) - \left(\mathbf{x}\boldsymbol{\beta}^{(t-1)}, \mathbf{x}\boldsymbol{\gamma}^{(t-1)}\right)\right) + ln\left(J_{((\xi_m,\upsilon_m)\to(s_m,p_m))}\right) \\ & + lnf_{\mathbf{x}}\left(\mathbf{x}|w_m, r_m, Q_m, \boldsymbol{\theta}_{\mathbf{x}}^{(t-1)}\right) \end{aligned}$$

I sample using the Random Walk Metropolis-Hastings algorithm. The details are as follows

Let $\mathbf{x}_m^{(t,k)}$ be the vector whose lth element is:

$$x_{m,l}^{(t,k)} = \begin{cases} x_{m,l}^{(t)} & \text{for } l < k \\ x_{m,l}^{(t-1)} & \text{for } l \geq k \end{cases}$$

and let $\mathbf{x}_m^{(t,k)\dagger}$ be the vector whose l th element is:

$$x_{m,l}^{(t,k)\dagger} = \begin{cases} x_{m,l}^{(t)} & \text{for } l < k \\ x_{m,l}^{(t-1)} + \epsilon_N,\ \epsilon_N \sim N\left(0, \tau_{xk}\right) & \text{for } l = k\ \cdot \\ x_{m,l}^{(t-1)} & \text{for } l > k \end{cases}$$

Then, draw uniform distribution $\epsilon_U \sim U[0,1]$, and let

$$x^{(t)}_{m,k} = \begin{cases} = \quad x^{(t,k)\dagger}_{m,k} & \text{if } \epsilon_U \leq \frac{exp\left(l_{\mathbf{x}}\left(s^{(t-1)}_m, p^{(t-1)}_m, w_m, r_m, Q_m, \mathbf{x}^{(t,k)\dagger}_m; \boldsymbol{\theta}^{(t-1)}\right)\right)}{exp\left(l_{\mathbf{x}}\left(s^{(t-1)}_m, p^{(t-1)}_m, w_m, r_m, Q_m, \mathbf{x}^{(t,k)}_m; \boldsymbol{\theta}^{(t-1)}\right)\right)} \\ = \quad x^{(t,k)}_{m,k} & \text{if otherwise} \end{cases}$$

2. Sampling (s,p) **for** $m \notin \mathcal{E}$: Given $\left(\mathbf{x}^{(t)}_m, w_m, r_m, Q_m; \boldsymbol{\theta}^{(t-1)}\right)$, I augment the missing (s_m, p_m) for each nonentrant firm $m \notin \mathcal{E}$. I do so by using the Metropolis-Hastings algorithm. Recall $h(\cdot)$ is specified as in Equations (40) and (41). Since the cost data is not observable, the log likelihood increment that includes (s_m, p_m) is $l_{hm} + l_{em}$, which is a function of (s,p) given $\left(w_m, r_m, Q_m, \mathbf{x}^{(t)}_m; \boldsymbol{\theta}^{(t-1)}\right)$ as below:

$$\begin{aligned} & l_{(s,p)}\left(\left(s,p\right), w_m, r_m, Q_m, \mathbf{x}^{(t)}_m; \boldsymbol{\theta}^{(t-1)}\right) \\ \equiv\ & lnf_{(\xi,\upsilon)}\left(h\left(\left(s,p\right)|w_m, r_m, Q_m, \mathbf{x}^{(t)}_m; \boldsymbol{\theta}^{(t-1)}\right) - \left(\mathbf{x}^{(t)}_m \boldsymbol{\beta}^{(t-1)}, \mathbf{x}^{(t)}_m \boldsymbol{\gamma}^{(t-1)}\right)\right) + ln\left(J_{((\xi_m,\upsilon_m)\rightarrow(s_m,p_m))}\right) \\ & -ln\left(1 + exp\left(pQ_m s - C^*_m\left(s, p, Q_m, \boldsymbol{\theta}^{(t-1)}\right) - E^{(t-1)}_C\right)\right),\ m \notin \mathcal{E}. \end{aligned}$$

We draw the proposal values (s_m, p_m) as follows:

$$\begin{aligned} s^{(t)\dagger}_m &= s^{(t-1)}_m + \epsilon_{z_1 m},\ \epsilon_{z_1 m} \sim N\left(0, \tau_1\right) \\ p^{(t)\dagger}_m &= p^{(t-1)}_m + \epsilon_{z_2 m},\ \ \epsilon_{z_2 m} \sim N\left(0, \tau_2\right). \end{aligned}$$

Then, draw uniform distribution $\epsilon_a \sim U[0,1]$, and let

$$\left(s^{(t)}_m, p^{(t)}_m\right) = \begin{cases} \left(s^{(t)\dagger}_m, p^{(t)\dagger}_m\right) & \text{if } \epsilon_a \leq \frac{exp\left(l_{(s,p)}\left(\left(s^{(t)\dagger}_m, p^{(t)\dagger}_m\right), w_m, r_m, Q_m, \mathbf{x}^{(t)}_m; \boldsymbol{\theta}^{(t-1)}\right)\right)}{exp\left(l_{(s,p)}\left(\left(s^{(t-1)}_m, p^{(t-1)}_m\right), w_m, r_m, Q_m, \mathbf{x}^{(t)}_m; \boldsymbol{\theta}^{(t-1)}\right)\right)} \\ \left(s^{(t-1)}_m, p^{(t-1)}_m\right) & \text{if otherwise} \end{cases}$$

4.3.2 Estimation

After sampling $\left(s^{(t)}_m, p^{(t)}_m, \mathbf{x}^{(t)}_m\right)$, I generate iteration t sample of the parameters. I first sample $\left(\alpha^{(t)}_c, \beta^{(t)}_c, \alpha^{(t)}, E^{(t)}_C\right)$ and then, the parameters of the distribution of the measurement errors. Below, I collect the terms of the log likelihood increment of firm m which contain the parameters

$(\alpha_c, \beta_c, \alpha, E_C)$, and construct the log likelihood as their sum over m.

$$\begin{aligned}
& l_{(\alpha_c,\beta_c,\alpha,E_C)m} \\
\equiv \; & \left[-\frac{K}{2}ln\pi - \frac{1}{2}ln\left|\boldsymbol{\Sigma}_{\mathbf{u}}\right| - \frac{1}{2}\mathbf{u}'_m\boldsymbol{\Sigma}_{\mathbf{u}}^{-1}\mathbf{u}_m\right] I\left(m \in \mathcal{E}\right) && (45) \\
& lnf_{(\xi,\upsilon)}\Big(h\left(\left(s_m, p_m\right)|w_m, r_m, \mathbf{x}_m\right) - \left(\mathbf{x}_m\boldsymbol{\beta}, \mathbf{x}_m\boldsymbol{\gamma}\right)\Big) + ln\left(J_{\left(\left(\xi_m,\upsilon_m\right)\rightarrow\left(s_m,p_m\right)\right)}\right) && (46) \\
& +ln\left(\frac{p_m q_m - C_m^* - E_C}{1 + exp\left(p_m q_m - C_m^* - E_C\right)}\right) I\left(m \in \mathcal{E}\right) \\
& +ln\left(\frac{1}{1 + exp\left(p_m q_m - C_m^* - E_C\right)}\right) I\left(m \notin \mathcal{E}\right). && (47) \\
& l_{(\alpha_c,\beta_c,\alpha,E_C)} = \sum_{m=1}^{M} l_{(\alpha_c,\beta_c,\alpha,E_C)m}
\end{aligned}$$

Since (α_c, β_c) are the parameters of the cost function, they enter in C_m^* through Equations (27). Since Equations (38) and (39) indicate $\mathbf{u}_m$ contains C_m^*, and so does (47), they all contain (α_c, β_c). Furthermore, from Equations (40) and (41), we can see that h and its Jacobian in (46) contain $(\alpha_c, \beta_c, \alpha)$. Furthermore, Equation (47) contains E_C.

Let $\mathbf{s}^{(t)}$ be the vector whose m^{th} element is $s_m^{(t)}$. Similarly for $\mathbf{p}^{(t)}$, $\mathbf{w}$, $\mathbf{r}$, $\mathbf{Q}$. Also, let $\mathbf{X}^{(t)}$, be the matrix whose m^{th} row is $\mathbf{x}_m^{(t)}$.

1. Sampling $\left(\alpha_c^{(t)}, \beta_c^{(t)}, \alpha^{(t)}, E_C^{(t)}\right)$: I use the Random -Walk Metropolis-Hastings algorithm to resample $\left(\alpha_c^{(t)}, \beta_c^{(t)}, \alpha^{(t)}, E_C^{(t)}\right)$. First, I resample $\alpha_c^{(t)}$ as follows. I generate the candicate $\alpha_c^{(t)\dagger} = \alpha_c^{(t-1)} + \epsilon_{\alpha_c}$, where $\epsilon_{\alpha_c} \sim N\left(0, \tau_{\alpha_c}\right)$. Then, draw $\epsilon_a \sim U\left[0, 1\right]$ and set

$$\alpha_c^{(t)} = \begin{cases} = \quad \alpha_c^{(t)\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\alpha_c^{(t)\dagger},\boldsymbol{\theta}^{(t-1)}\backslash\alpha_c^{(t-1)}\right)\right)}{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\boldsymbol{\theta}^{(t-1)}\right)\right)} \\ = \quad \alpha_c^{(t-1)} & \text{if otherwise.} \end{cases}$$

I denote $\widetilde{\boldsymbol{\theta}} \equiv \left(\boldsymbol{\theta}^{(t-1)} \setminus \alpha_c^{(t-1)}, \alpha_c^{(t)}\right)$ and I do the same procedure for β_c. I generate the candicate $\beta_c^{(t)\dagger} = \beta_c^{(t-1)} + \epsilon_{\beta_c}$, where $\epsilon_{\beta_c} \sim N\left(0, \tau_{\beta_c}\right)$. Then, I draw $\epsilon_a \sim U\left[0, 1\right]$ and set

$$\beta_c^{(t)} = \begin{cases} = \quad \beta_c^{(t)\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\beta_c^{(t)\dagger},\widetilde{\boldsymbol{\theta}}\backslash\beta_c^{(t-1)}\right)\right)}{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},,\mathcal{E};\widetilde{\boldsymbol{\theta}}\right)\right)} \\ = \quad \beta_c^{(t-1)} & \text{if otherwise.} \end{cases}$$

Next, I draw $\alpha^{(t)}$ in the same way, by using the Metropolis-Hastings algorithm. I again use the random-walk Metropolis-Hastings algorithm. I denote $\widetilde{\boldsymbol{\theta}} \equiv \left(\boldsymbol{\theta}^{(t-1)} \setminus \left(\alpha_c^{(t-1)}, \beta_c^{(t-1)}\right), \left(\alpha_c^{(t)}, \beta_c^{(t)}\right)\right)$. As before, I generate the candidate $\alpha^{(t)\dagger} = \alpha^{(t-1)} + \epsilon_\alpha$, where $\epsilon_\alpha \sim N\left(0, \tau_\alpha\right)$. Then, draw

$\epsilon_a \sim U[0,1]$ and

$$\alpha^{(t)} = \begin{cases} = & \alpha^{(t)\dagger} \quad \text{if } \epsilon_a \leq \frac{exp\left(l\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},,\mathcal{E};\alpha^{(t)\dagger},\widetilde{\boldsymbol{\theta}}\backslash\alpha^{(t-1)}\right)\right)}{exp\left(l\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},,\mathcal{E};\widetilde{\boldsymbol{\theta}}\right)\right)} \\ = & \alpha^{(t-1)} \quad \text{if otherwise.} \end{cases}$$

Next, I draw $E_C^{(t)}$ in the same way, by using the Metropolis-Hastings algorithm. I again use the random-walk Metropolis-Hastings algorithm. I denote $\widetilde{\boldsymbol{\theta}} \equiv \left(\boldsymbol{\theta}^{(t-1)} \setminus \left(\alpha_c^{(t-1)}, \beta_c^{(t-1)}, \alpha^{(t-1)}\right), \left(\alpha_c^{(t)}, \beta_c^{(t)}, \alpha^{(t)}\right)\right)$. As before, I generate the candidate $E_C^{(t)\dagger} = E_C^{(t-1)} + \epsilon_{E_C}$, where $\epsilon_{E_C} \sim N(0, \tau_{E_C})$. Then, draw $\epsilon_{E_C} \sim U[0,1]$ and

$$E_C^{(t)} = \begin{cases} = & E_C^{(t)\dagger} \quad \text{if } \epsilon_{E_C} \leq \frac{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},,\mathcal{E};E_C^{(t)\dagger},\widetilde{\boldsymbol{\theta}}\backslash E_C^{(t-1)}\right)\right)}{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\widetilde{\boldsymbol{\theta}}\right)\right)} \\ = & E_C^{(t-1)} \quad \text{if otherwise.} \end{cases}$$

2. Sampling $\left(\boldsymbol{\beta}^{(t)}, \sigma_{\boldsymbol{\beta}}^{(t)}\right)$, $\left(\boldsymbol{\gamma}^{(t)}, \sigma_{\boldsymbol{\gamma}}^{(t)}\right)$: Note that $(\boldsymbol{\beta}, \boldsymbol{\gamma})$ only enter the log likelihood in Equation (44) in $lnf_{(\xi,\upsilon)}\left(h\left((s_m,p_m)|w_m,r_m,Q_m\right) - (\mathbf{x}_m\boldsymbol{\beta}, \mathbf{x}_m\boldsymbol{\gamma})\right)$ in the following form:

$$\begin{aligned} \delta_m &= h_1\left((s_m,p_m)|w_m,r_m,Q_m\right) = \mathbf{x}_m\boldsymbol{\beta} + \xi_m \\ \eta_m &= h_2\left((s_m,p_m)|w_m,r_m,Q_m\right) = \mathbf{x}_m\boldsymbol{\gamma} + \upsilon_m \end{aligned}$$

Therefore, we sample $\boldsymbol{\beta}^{(t)}$ using the below linear regression equation with $\boldsymbol{\beta}$ to be the unknown parameter to be estimated.

$$\delta_m^{(t)} = lns_m^{(t)} - ln\left(1 - s_m^{(t)}\right) + p_m^{(t)}\alpha^{(t)} = \mathbf{x}_m^{(t)}\boldsymbol{\beta} + \xi_m.$$

Let $\boldsymbol{\delta}^{(t)}$ be the vector whose m^{th} element is $\delta_m^{(t)}$. Then, $\boldsymbol{\beta}^{(t)} \sim N\left(\widehat{\boldsymbol{\beta}}, \sigma_{\boldsymbol{\beta}}^{(t-1)2}\left(\mathbf{X}^{(t)\prime}\mathbf{X}^{(t)}\right)^{-1}\right)$, where

$$\widehat{\boldsymbol{\beta}} = \left[\mathbf{X}^{(t)\prime}\mathbf{X}^{(t)}\right]^{-1}\mathbf{X}^{(t)\prime}\boldsymbol{\delta}^{(t)}.$$

Then, draw $\sigma_{\boldsymbol{\beta}}^{(t)}$ as follows:

$$\sigma_{\boldsymbol{\beta}}^{(t)2} \sim IG\left(\frac{M}{2}, \frac{\left(\boldsymbol{\delta}^{(t)} - \mathbf{X}^{(t)}\boldsymbol{\beta}^{(t)}\right)'\left(\boldsymbol{\delta}^{(t)} - \mathbf{X}^{(t)}\boldsymbol{\beta}^{(t)}\right)}{2}\right).$$

Next, I sample $\left(\boldsymbol{\gamma}^{(t)}, \sigma_{\boldsymbol{\gamma}}^{(t)}\right)$. I can use Equation (41) to derive $\eta_m^{(t)}$ as follows.

$$\begin{aligned}\eta_m^{(t)} &= ln\left(p_m^{(t)} - \frac{1}{\left(1-s_m^{(t)}\right)\alpha^{(t)}}\right)\\ &-\Big[\frac{\alpha_c^{(t)}}{\alpha_c^{(t)}+\beta_c^{(t)}} ln\left(\frac{w_m}{\alpha_c^{(t)}}\right) + \frac{\beta_c^{(t)}}{\alpha_c^{(t)}+\beta_c^{(t)}} ln\left(\frac{r_m}{\beta_c^{(t)}}\right)\\ &+\left(\frac{1}{\alpha_c^{(t)}+\beta_c^{(t)}} - 1\right)\left(lnQ_m + lns_m^{(t)}\right)\Big] = \mathbf{x}_m^{(t)}\boldsymbol{\gamma} + \upsilon_m \end{aligned} \tag{48}$$

let $\boldsymbol{\eta}^{(t)}$ be the vector whose m^{th} element is $\eta_m^{(t)}$. Then, $\boldsymbol{\gamma}^{(t)} \sim N\left(\widehat{\boldsymbol{\gamma}}, \sigma_{\boldsymbol{\gamma}}^{(t-1)2}\left(\mathbf{X}^{(t)\prime}\mathbf{X}^{(t)}\right)^{-1}\right)$, where

$$\widehat{\boldsymbol{\gamma}} = \left[\mathbf{X}^{(t)\prime}\mathbf{X}^{(t)}\right]^{-1}\mathbf{X}^{(t)\prime}\boldsymbol{\eta}^{(t)}$$

Then, draw $\sigma_{\boldsymbol{\gamma}}^{(t)}$ as follows:

$$\sigma_{\boldsymbol{\gamma}}^{(t)2} \sim IG\left(\frac{M}{2}, \frac{\left(\boldsymbol{\eta}^{(t)} - \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}\right)'\left(\boldsymbol{\eta}^{(t)} - \mathbf{X}^{(t)}\boldsymbol{\gamma}^{(t)}\right)}{2}\right)$$

Finally, I draw $\boldsymbol{\Sigma}_{\mathbf{u}}^{(t)}$ as inverse Wishart distribution as follows.

$$\boldsymbol{\Sigma}_{\mathbf{u}}^{(t)-1} = G \sim W\left(M_{\mathcal{E}}, \sum_{m=1}^{M_{\mathcal{E}}} \mathbf{u}_m^{(t)}\mathbf{u}_m^{(t)\prime}\right)$$

where $M_{\mathcal{E}}$ is the number of entrant firms.

I report the results of the Monte-Carlo simulation using the method described above. The initial parameters used here are identical to those in Table 1. I generated 4,000 hypothetical markets where each market has one firm that either enters as a monopolist or remains a potential entrant. I iterated the MCMC sampling process 10,000 times. I designated the first 5,000 samples as the burn-in samples and only used the remaining samples for the posterior distribution. Table 5 displays the mean and standard deviation of the posterior distributions of the parameters. We can see that the means are close to their true values, and the standard deviations are relatively small. Figures 5 and 6 plot the generated parameters during the MCMC sampling process. The figures clearly show the convergence of the MCMC sampled parameters to the true parameters. These findings suggest that the estimation method successfully recovered the true parameters.

4.4 Estimation of monopoly entry without cost data.

Next, I present the results of the Monte-Carlo study where I modify the above MCMC sampling procedures and assume that data on cost is not observable. Since the observed data does not include data on cost, I need to eliminate the components of the likelihood function that includes cost data as variable. That is, we eliminate l_{um} from the log likelihood function in Equation (44). Furthermore, in order to better relate to the identification argument, we express the profit function as in Equation (34). That is,

$$\begin{aligned}\pi_m &= p_m q_m - C_m^* = q_m \left[p_m - (\alpha_c + \beta_c) \left(p_m - \frac{1}{(1 - s_m)\,\alpha} \right) \right] \\ &= q_m \left[(1 - \alpha_c - \beta_c)\, p_m + \frac{\alpha_c + \beta_c}{\alpha} \frac{1}{(1 - s_m)} \right]\end{aligned}$$

I then have the following modified likelihood:

$$\begin{aligned} & l_m \left(s_m, p_m, w_m, r_m, Q_m, \mathcal{E} \right) \\ \equiv\ & l_{hm} + l_{em} + l_{\mathbf{x}m} \\ =\ & lnf_{(\xi,\upsilon)} \left(h \left((s_m, p_m) \,| w_m, r_m, Q_m \right) - (\mathbf{x}_m \boldsymbol{\beta}, \mathbf{x}_m \boldsymbol{\gamma}) \right) + ln \left(J_{((\xi_m, \upsilon_m) \to (s_m, p_m))} \right) \\ & + ln \left(\frac{\pi_m - E_C}{1 + exp\left(\pi_m - E_C \right)} \right) I \left(m \in \mathcal{E} \right) + ln \left(\frac{1}{1 + exp\left(\pi_m - E_C \right)} \right) I \left(m \notin \mathcal{E} \right). \\ & + lnf_{\mathbf{x}} \left(\mathbf{x}_m | w_m, r_m, Q_m, \boldsymbol{\theta}_{\mathbf{x}} \right) \end{aligned} \tag{49}$$

where

$$\pi_m = q_m \left[(1 - \alpha_c - \beta_c)\, p_m + \frac{\alpha_c + \beta_c}{\alpha} \frac{1}{(1 - s_m)} \right] \tag{50}$$

Data Augmentation: Let us now recall the data augmentation algorithm for the estimation when cost data was available. Note that even in this case, for nonentrants, cost data was not available. Hence, I could not construct the likelihood increment l_{um} for them. That is why I only used the likelihood increments l_{hm}, l_{em}, and $l_{\mathbf{x}m}$ for the data augmentation of $(s_m, p_m, \mathbf{x}_m)$. Because the likelihood increments l_{hm}, l_{em}, and $l_{\mathbf{x}m}$ can be constructed even if the cost data is not available, which is the case here, there is no need for any modification of the data augmentation procedure. That is, I will conduct the data augmentation algorithm for nonentrant firms exactly the same as before when the cost data were available for entrants.

Estimation: Through data augmentation, I generated missing $\left(s_m^{(t)}, p_m^{(t)}, \mathbf{x}_m^{(t)} \right)$ for the non-entrant firms. I now estimate the subset of the demand, cost, and entry cost parameters

$\left(\alpha_c^{(t)}, \beta_c^{(t)}, \alpha^{(t)}, E_C^{(t)}\right)$ and the parameters of the distribution of the measurement errors.

$$\begin{aligned} & l_{(\alpha_c,\beta_c,\alpha,E_C)m}\left(s_m^{(t)}, p_m^{(t)}, w_m, r_m, \mathbf{x}_m^{(t)}, Q_m, \mathcal{E}; \boldsymbol{\theta}^{(t-1)}\right) && (51) \\ = \; & lnf_{(\xi,\upsilon)}\left(h\left(\left(s_m^{(t)}, p_m^{(t)}\right)|w_m, r_m, Q_m\right) - \left(\mathbf{x}_m^{(t)}\boldsymbol{\beta}^{(t-1)}, \mathbf{x}_m^{(t)}\boldsymbol{\gamma}^{(t-1)}\right)\right) \\ & + ln\left(J_{((\xi_m,\upsilon_m)\to(s_m,p_m))}\right) && (52) \\ & + ln\left(\frac{exp\left(\pi_m - E_C^{(t-1)}\right)}{1+exp\left(\pi_m - E_C^{(t-1)}\right)}\right) I\left(m \in \mathcal{E}\right) \\ & + ln\left(\frac{1}{1+exp\left(\pi_m - E_C^{(t-1)}\right)}\right) I\left(m \notin \mathcal{E}\right). && (53) \\ & \pi_m = s_m^{(t)} Q_m\left[\left(1-\alpha_c^{(t-1)}-\beta_c^{(t-1)}\right)p_m^{(t)} + \frac{\alpha_c^{(t-1)}+\beta_c^{(t-1)}}{\alpha^{(t-1)}}\frac{1}{\left(1-s_m^{(t)}\right)}\right] && (54) \\ & l_{(\alpha_c,\beta_c,\alpha,E_C)} = \sum_{m=1}^{M} l_{(\alpha_c,\beta_c,\alpha,E_C)m} && (55) \end{aligned}$$

As before, I use the Random-Walk Metropolis-Hastings algorithm to resample α_c as follows. I generate the candidate $\alpha_c^{(t)\dagger} = \alpha_c^{(t-1)} + \epsilon_{\alpha_c}$, where $\epsilon_{\alpha_c} \sim N\left(0, \tau_{\alpha_c}\right)$. Then, draw $\epsilon_a \sim U\left[0,1\right]$ and

$$\alpha_c^{(t)} = \begin{cases} = \quad \alpha_c^{(t)\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\alpha_c^{(t)\dagger},\boldsymbol{\theta}^{(t-1)}\setminus\alpha_c^{(t-1)}\right)\right)}{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\boldsymbol{\theta}^{(t-1)}\right)\right)} \\ = \quad \alpha_c^{(t-1)} & \text{if otherwise} \end{cases}$$

I denote $\widetilde{\boldsymbol{\theta}} \equiv \left(\boldsymbol{\theta}^{(t-1)} \setminus \alpha_c^{(t-1)}, \alpha_c^{(t)}\right)$. Similarly, I use the Random -Walk Metropolis-Hastings algorithm to sample $\beta_c^{(t)}$ as follows. I generate the candidate $\beta_c^{(t)\dagger} = \beta_c^{(t-1)} + \epsilon_{\beta_c}$, where $\epsilon_{\beta_c} \sim N\left(0, \tau_{\beta_c}\right)$. Then, draw $\epsilon_a \sim U\left[0,1\right]$ and

$$\beta_c^{(t)} = \begin{cases} = \quad \beta_c^{(t)\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\beta_c^{(t)\dagger},\widetilde{\boldsymbol{\theta}}\setminus\beta_c^{(t-1)}\right)\right)}{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\widetilde{\boldsymbol{\theta}}\right)\right)} \\ = \quad \beta_c^{(t-1)} & \text{if otherwise} \end{cases}$$

Next, I draw α in the same way, by using the Metropolis-Hastings algorithm. I denote $\widetilde{\boldsymbol{\theta}} \equiv \left(\boldsymbol{\theta}^{(t-1)} \setminus \left(\alpha_c^{(t-1)}, \beta_c^{(t-1)}\right), \left(\alpha_c^{(t)}, \beta_c^{(t)}\right)\right)$. I again use the random-walk Metropolis-Hastings algorithm. As before, I generate the candidate $\alpha^{(t)\dagger} = \alpha^{(t)} + \epsilon_\alpha$, where $\epsilon_\alpha \sim N\left(0, \tau_\alpha\right)$. Then, draw $\epsilon_a \sim U[0,1]$ and

$$\alpha^{(t)} = \begin{cases} = \quad \alpha^{(t)\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\alpha^{(t)\dagger},\widetilde{\boldsymbol{\theta}}\backslash\alpha^{(t-1)}\right)\right)}{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\widetilde{\boldsymbol{\theta}}\right)\right)} \\ = \quad \alpha^{(t-1)} & \text{if otherwise} \end{cases}$$

Next, I sample $E_C^{(t)}$ in the same way. I denote $\widetilde{\boldsymbol{\theta}} \equiv \left(\boldsymbol{\theta}^{(t-1)} \setminus \left(\alpha_c^{(t-1)}, \beta_c^{(t-1)}, \alpha^{(t-1)}\right), \left(\alpha_c^{(t)}, \beta_c^{(t)}, \alpha^{(t)}\right)\right)$. I again use the random-walk Metropolis-Hastings algorithm. As before, I generate the candidate $E_C^{(t)\dagger} = E_C^{(t)} + \epsilon_{E_C}$, where $\epsilon_{E_C} \sim N(0, \tau_{E_C})$. Then, draw $\epsilon_{E_C} \sim U[0,1]$ and

$$E_C^{(t)} = \begin{cases} = \quad E_C^{(t)\dagger} & \text{if } \epsilon_{E_C} \leq \frac{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};E_C^{(t)\dagger},\widetilde{\boldsymbol{\theta}}\backslash E_C^{(t-1)}\right)\right)}{exp\left(l_{(\alpha_c,\beta_c,\alpha,E_C)}\left(\mathbf{s}^{(t)},\mathbf{p}^{(t)},\mathbf{w},\mathbf{r},\mathbf{X}^{(t)},\mathbf{Q},\mathcal{E};\widetilde{\boldsymbol{\theta}}\right)\right)} \\ = \quad E_C^{(t-1)} & \text{if otherwise} \end{cases}$$

2. Sampling $\left(\boldsymbol{\beta}^{(t)}, \sigma_{\boldsymbol{\beta}}^{(t)}\right)$, $\left(\boldsymbol{\gamma}^{(t)}, \sigma_{\boldsymbol{\gamma}}^{(t)}\right)$: I do the sampling in exactly the same way as in the case of monopoly with cost data. Hence, I omit the exposition.

I conduct two Monte-Carlo simulations to verify the accuracy of the estimation method mentioned in this subsection. I describe the results in Tables 6 and 7. The results in the two tables are obtained from the estimation of the model generated samples under two different parameter configurations. In particular, Table 6 presents the estimation of the model with a relatively smaller variance for the variables (x, ξ, υ) compared to the results in Table 7. Similar to the analysis conducted in Table 5, I generated 4,000 hypothetical markets and ran the MCMC simulation 10,000 times.

The results indicate that the estimates of the structural parameters are accurate when the variance of the exogenous variables is small. However, the estimates tend to be biased as the variance increases. In particular, the posterior means α_c, β_c are 0.436 and 0.315, respectively. Bias in α_c suggests failure of identification based on Equation (93). This underscores the importance of including both demand and cost data in the estimation of structural parameters, rather than relying solely on demand data from new entrants, especially when the sample size is small.

5 Oligopoly Model

In the earlier sections, I simplified the discussion by assuming only one potential entrant per market, sidestepping the complexity of the computation of the entry equilibria. Now, I am broadening the scope to include an oligopoly model. In order to avoid having to deal with the issue of multiple equilibria pointed out by Ciliberto and Tamer (2009), Ciliberto et al. (2021) and others, I introduce the concept of sequential entry. In this extended model, I assume that nature assigns the probability of the timing of each firm's move, which I also assume that the econometrician has the knowledge of.[10] Even though this is a strong assumption, it is useful because it is the key assumption to ensuring that the equilibrium outcomes are unique. The methodology for identifying these outcomes remains consistent with that of the monopoly model. I also showcase the results of Monte-Carlo simulations to validate my approach. These simulations demonstrate that my method can accurately and reliably estimate the parameters in question.

5.1 Random Timing of Entry Decision

Assume an oligopoly model where firms decide whether to enter the market or stay out in period 1 and when they enter, compete for prices in period 2. Let $\mathbf{e}_m$ be the vector of each firm's choice in market m, and e_{jm} is the choice of firm j. Firm j in market m chooses whether to enter or stay out. That is,

$$e_{jm} = \begin{cases} 1 & if\ enter \\ 0 & otherwise \end{cases}$$

As before, I denote $\boldsymbol{\xi}_m$ to be the $J_m \times 1$ vector whose jth elemct is ξ_{jm}, the unobserved product characteristics of firm $j \in \{1, \dots J_m\}$ in market $m \in \{1, \dots, M\}$. $\boldsymbol{v}_m$, $\mathbf{s}_m$, $\mathbf{p}_m, \mathbf{q}_m$ are similarly defined. Similarly, bold characters represent vectors, except for $\mathbf{X}_m = (\mathbf{x}_{1m}, \mathbf{x}_{2m}, \dots, \mathbf{x}_{jm}, \dots, \mathbf{x}_{J_m m})$, which is a product characteristics matrix consisting of vectors of characteristics of firm j in market m. Let us assume that in period 1 firms choose the entry/exit decisions sequentially. Firm j becomes the kth one to decide on entry with probability κ_{jkm}. This probability is assumed to be observed by the econometrician. If the econometrician does not have this information, it is natural to assume that all firms have equal probability to be kth firm to decide, that is, $\kappa_{jkm} = \frac{1}{J_m}$ for all j, k, given an m.

5.1.1 Period 2: Price Competition

Let us first consider the period 2 where only n out of J firms entered the market. In this example, we use the same market share function and the cost function in the monopoly model. Then,

[10] The idea of assigning orders of entry is not new. For example, Li et al. (2022)assumes that firms that have higher presence in each airport can move earlier to avoid the multiplicity of equilibria and estimate the demand system.

as shown by Mizuno (2003), it is known in the literature that the price competition in such an oligopoly model, given the entrant firms, is unique. Then, taking the price of other entrant firms as given, each firm chooses the price according to the following profit maximizing problem:

$$arg \max_{p_{jm}} \ \pi_{jm}(\boldsymbol{p}_m, \mathbf{X}_m, \mathbf{w}_m, \boldsymbol{\xi}_{jm}, \upsilon_{jm}, Q_m, \theta)$$

where

$$\pi_{jm} \equiv p_{jm} \times Q_m \times s_{jm}\left(\boldsymbol{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_{jm}; \theta_d\right) - C_{jm}(s_{jm}\left(\boldsymbol{p}_m, \mathbf{X}_m, \boldsymbol{\xi}_m; \theta_d\right), Q_m, \mathbf{w}_m, \upsilon_{jm}, \theta_c)$$

Let us denote the equilibrium price of firm j in market m under n entrants as p^n_{jm}, and their profit as π^n_{jm}.

5.1.2 Period 1: Entry Decision

Let us assume that before entry, firms draw the random profit shock of entry ϵ^E_{jm} and the shock of not entering the market ϵ_{jm}. Firm j with n firms in the market will enter the market as long as $\pi^n_{jm} - E_C + \epsilon^E_{jm} > \epsilon_{jm}$, where E_C is the entry cost. Let us write the net benefit from entry, including the entry as: $\pi^{n*}_{jm} \equiv \pi^n_{jm} - E_C + \epsilon^E_{jm} - \epsilon_{jm}$.

5.1.3 Entry Equilibria

Firms move sequentially according to the probability of order the nature assigned them. Once the order of the movement is realized, the game of firms' sequential entry has a unique equilibrium. As an illustrative example, let us consider an entry game where there are two potential entrants in market m and that firm 1 is randomly assigned to move first. Denote $\mathbf{e}_m \equiv (e_{1m}, e_{2m})$ to be the vector of entry decision, where e_{jm} equals one if firm j chooses to enter and zero otherwise. Then, there are four possible outcomes: $\mathbf{e}_m \in \{(1,1), (1,0), (0,1), (0,0)\}$. Those equilibria are realized in the following scenario:

1. If $\pi^{2*}_{1m} > 0, \pi^{2*}_{2m} > 0$, then $\mathbf{e}_m = (1,1)$ is the Nash equilibrium.

2. If $\pi^{2*}_{1m} > 0, \pi^{2*}_{2m} \leq 0$ or $\pi^{2*}_{1m} \leq 0, \pi^{2*}_{2m} \leq 0, \pi^{1*}_{1m} > 0$, then $\mathbf{e}_m = (1,0)$ is the Nash equilibrium.

3. If $\pi^{2*}_{1m} \leq 0, \pi^{2*}_{2m} > 0$ or $\pi^{2*}_{1m} \leq 0, \pi^{2*}_{2m} \leq 0, \pi^{1*}_{1m} \leq 0, \pi^{1*}_{2m} > 0$, then $\mathbf{e}_m = (0,1)$ is the Nash equilibrium.

4. If $\pi^{1*}_{1m} \leq 0, \pi^{1*}_{2m} \leq 0$, then $\mathbf{e}_m = (0,0)$ is the Nash equilibrium.

5.1.4 Entry Probability

The entry probability is the the probability of realizing each Nash equilibrium. For example, if there are two potential entrants in the market and the firm 1 moves first with a probability of 50%, the entry probability is given as follows: First, denote

$$\mathbf{Z}_m \equiv (\mathbf{X}_m, \boldsymbol{\delta}_m, \boldsymbol{\eta}_m, \mathbf{w}_m, Q_m)$$

to be the vectors observed and unobserved product characteristics, cost shock for all firms, and the vector of input prices and the market size of market m. Given those variables in market m, and given the entry behavior of rival firms, firm j derives her profit from the monopoly/oligopoly equilibrium of the pricing game. Below, I denote $\mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m = (1,1))$ to be the set of entry and outside option shocks $\left(\boldsymbol{\epsilon}_m^E, \boldsymbol{\epsilon}_m^O\right)$ that results in the equilibrium entry vector to be $(1,1)$. I denote the same way for $\mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m = (0,0))$. Note, that the set $\mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m = (1,0))$ is different depending on whether the firm 1 is the first mover or not. Hence, I additionally denote (l, l') in which l is the first mover firm and l' is the 2nd mover firm. Then,

$$\begin{aligned}
\mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m = (1,1)) &= \left\{\left(\boldsymbol{\epsilon}_m^E, \boldsymbol{\epsilon}_m^O\right) : \pi_{1m}^{2*} > 0, \pi_{2m}^{2*} > 0\right\} \\
\mathcal{A}\left(\mathbf{Z}_m, \mathbf{e}_m = (1,0), \left(l, l'\right) = (1,2)\right) &= \left\{\left(\boldsymbol{\epsilon}_m^E, \boldsymbol{\epsilon}_m^O\right) : \left(\pi_{lm}^{2*} > 0, \pi_{l'm}^{2*} \leq 0\right)\right. \\
&\quad \left.\cup \left(\pi_{lm}^{2*} \leq 0, \pi_{l'm}^{2*} \leq 0, \pi_{lm}^{1*} > 0\right)\right\} \\
\mathcal{A}\left(\mathbf{Z}_m, \mathbf{e}_m = (0,1), \left(l, l'\right) = (1,2)\right) &= \left\{\left(\boldsymbol{\epsilon}_m^E, \boldsymbol{\epsilon}_m^O\right) : \left(\pi_{lm}^{2*} \leq 0, \pi_{l'm}^{2*} > 0\right)\right. \\
&\quad \left.\cup \left(\pi_{lm}^{2*} \leq 0, \pi_{l'm}^{2*} \leq 0, \pi_{lm}^{1*} \leq 0, \pi_{l'm}^{1*} > 0\right)\right\} \\
\mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m = (0,0)) &= \left\{\left(\boldsymbol{\epsilon}_m^E, \boldsymbol{\epsilon}_m^O\right) : \pi_{1m}^{1*} \leq 0, \pi_{2m}^{1*} \leq 0\right\}.
\end{aligned} \tag{56}$$

Then, the entry probability for $\mathbf{e}_m \in \{(0,0), (1,1)\}$ conditional on $\mathbf{Z}_m$ is

$$Pr(\mathbf{e}_m | \mathbf{Z}_m) = \int_{\boldsymbol{\epsilon} \in \mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m)} f(\boldsymbol{\epsilon}) \, d\boldsymbol{\epsilon}$$

and for $\mathbf{e}_m \in \{(0,1), (1,0)\}$

$$Pr(\mathbf{e}_m | \mathbf{Z}_m) = \frac{1}{2} \int_{\boldsymbol{\epsilon} \in \mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m, (l,l')=(1,2))} f(\boldsymbol{\epsilon}) \, d\boldsymbol{\epsilon} + \frac{1}{2} \int_{\boldsymbol{\epsilon} \in \mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m, (l,l')=(2,1))} f(\boldsymbol{\epsilon}) \, d\boldsymbol{\epsilon}$$

5.2 Likelihood Function

I carry the same definition for $\delta_{jm} = \mathbf{x}_{jm}\boldsymbol{\beta} + \xi_{jm}$ and $\eta_{jm} = \mathbf{x}_{jm}\boldsymbol{\gamma} + \upsilon_{jm}$ as section 4 and define $(\boldsymbol{\delta}_m, \boldsymbol{\eta}_m)$ as the vector of (δ_{jm}, η_{jm}) in market m.

Then, in the market where $\mathbf{e}_m = (1,1)$, (s_{jm}, p_{jm}), $j = 1,2$ are observable outcome of a duopoly game. Then, $(\boldsymbol{\delta}_m, \boldsymbol{\eta}_m)$ can be recovered by inverting the duopoly model as follows:

$$(\boldsymbol{\delta}_m, \boldsymbol{\eta}_m) = h\left((\boldsymbol{s}_m, \boldsymbol{p}_m) | w_m, r_m, Q_m\right).$$

In other markets, at least one of the firm is not an entrant, thus, its market share and price are unobservable. In that case, in constructing the likelihood, I assume (s_{jm}, p_{jm}) $j = 1,2$ when each firm is a monopolist is observable. Then, as in the monopolist case, (δ_{jm}, η_{jm}) can be recovered as follows:

$$(\delta_{jm}, \eta_{jm}) = h\left((s_{jm}, p_{jm}) | w_m, r_m, Q_m\right) \;\; j = 1,2.$$

We will see later that for each market, we will generate such (s_{jm}, p_{jm}) $j = 1,2$. Then, the log likelihood increment of market m is

$$\begin{aligned} & l\left(\mathbf{X}_m, \boldsymbol{\delta}_m, \boldsymbol{\eta}_m, \boldsymbol{e}_m | \mathbf{w}_m, Q_m\right) \\ = \; & lnf\left(\mathbf{X}_m | \mathbf{w}_m, Q_m\right) + lnf_{(\boldsymbol{\xi},\boldsymbol{\upsilon})}\left(\boldsymbol{\delta}_m - \mathbf{X}_m\boldsymbol{\beta}, \boldsymbol{\eta}_m - \mathbf{X}_m\boldsymbol{\gamma}\right) \\ & + ln\left(J\left(\mathbf{e}_m\right)\right) + lnPr\left(\mathbf{e}_m | \mathbf{Z}_m\right), \end{aligned}$$

where

$$ln\left(J\left(\mathbf{e}_m\right)\right) = \begin{cases} \sum_{j=1}^{2} ln\left\|\frac{\partial(\xi_{jm}, \upsilon_{jm})}{\partial(s_{jm}, p_{jm})}\right\| & if\ \mathbf{e}_m \in \{(0,0), (0,1), (1,0)\} \\ ln\left\|\frac{\partial(\boldsymbol{\xi}_m, \boldsymbol{\upsilon}_m)}{\partial(\mathbf{s}_m, \mathbf{p}_m)}\right\| & if\ otherwise \end{cases}$$

$$Pr\left(\mathbf{e}_m | \mathbf{Z}_m\right) = \begin{cases} \int_{\boldsymbol{\epsilon} \in \mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m)} f\left(\boldsymbol{\epsilon}\right) d\boldsymbol{\epsilon} & if\ \mathbf{e}_m \in \{(0,0), (1,1)\} \\ \frac{1}{2}\int_{\boldsymbol{\epsilon} \in \mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m, (l,l')=(1,2))} f\left(\boldsymbol{\epsilon}\right) d\epsilon & \\ +\frac{1}{2}\int_{\boldsymbol{\epsilon} \in \mathcal{A}(\mathbf{Z}_m, \mathbf{e}_m, (l,l')=(2,1))} f\left(\boldsymbol{\epsilon}\right) d\epsilon & if\ otherwise \end{cases}$$

Denote $\mathbf{H} \equiv (\mathbf{s}_m, \mathbf{p}_m, w_m, r_m, Q_m)$. Then, the first component of the log likelihood is

$$\begin{aligned} l_{hm} &= ln\Big[f(\mathbf{X}_m|\mathbf{w}_m, Q_m)\, f_{(\boldsymbol{\xi},\boldsymbol{\upsilon})}(\boldsymbol{\delta}_m(\mathbf{H}_m) - \mathbf{X}_m\boldsymbol{\beta}, \boldsymbol{\eta}_m(\mathbf{H}_m) - \mathbf{X}_m\boldsymbol{\gamma}) \\ &\times J(\mathbf{e}_m)\int_{\boldsymbol{\epsilon}\in\mathcal{A}(\mathbf{Z}_m,\mathbf{e}_m)} f(\boldsymbol{\epsilon})\, d\boldsymbol{\epsilon}\Big] \\ &= lnf(\mathbf{X}_m|\mathbf{w}_m, Q_m) + lnf_{(\boldsymbol{\xi},\boldsymbol{\upsilon})}(\boldsymbol{\delta}_m(\mathbf{H}_m) - \mathbf{X}_m\boldsymbol{\beta}, \boldsymbol{\eta}_m(\mathbf{H}_m) - \mathbf{X}_m\boldsymbol{\gamma}) + ln(J(\mathbf{e}_m)) \\ &+ lnPr(\mathbf{e}_m|\mathbf{Z}_m) \end{aligned}$$

As the second component of the log-likelihood function, I add the following likelihood derived by (36) to identify $(\alpha, \alpha_c, \beta_c)$:

$$l_{um} = \sum_{j}^{J_m}\left(-ln\pi - \frac{1}{2}ln|\boldsymbol{\Sigma}_{\mathbf{u}}| - \frac{1}{2}\mathbf{u}'_{jm}\boldsymbol{\Sigma}_{\mathbf{u}}^{-1}\mathbf{u}_{jm}\right)\cdot I(mj \in \mathcal{E}) \tag{57}$$

which the definitions of the variables inside (57) is identical to (36).

Using the two components, I define the likelihood function for market m as:

$$\begin{aligned} l_m = l_{hm} + l_{um} &= \sum_{j}^{J_m}\left(-ln\pi - \frac{1}{2}ln|\boldsymbol{\Sigma}_{\mathbf{u}}| - \frac{1}{2}\mathbf{u}'_{jm}\boldsymbol{\Sigma}_{\mathbf{u}}^{-1}\mathbf{u}_{jm}\right)\cdot I(mj \in \mathcal{E}) \\ &+ lnf(\mathbf{X}_m|w_m, r_m, Q_m) + lnf_{(\boldsymbol{\xi},\boldsymbol{\upsilon})}(\boldsymbol{\delta}_m(\mathbf{H}_m) - \mathbf{X}_m\boldsymbol{\beta}, \boldsymbol{\eta}_m(\mathbf{H}_m) - \mathbf{X}_m\boldsymbol{\gamma}) \\ &+ ln(J(\mathbf{e}_m)) + lnPr(\mathbf{e}_m|\mathbf{Z}_m) \end{aligned} \tag{58}$$

5.3 Estimation Procedure

This section describes the estimation of the structural parameters.

5.3.1 Data Augmentation

I simulate the variables product characteristics, demand shocks, and cost shocks for nonentrant firms using the above likelihood function. The augmentation technique is identical to the section 4, but this time I augment the variables for all nonentrants in the market.

First, I augment the product characteristics using the M-H algorithm.

1. Sampling $\mathbf{x}_{jm}$ for $jm \notin \mathcal{E}$: I first augment $\mathbf{x}_{jm}^{(t)}$. To do so, I use the Random walk Metropolis-Hastings algorithm, where, from $k = 1$ to K of unobservables, I successively draw the proposal value for the kth element of the vector $\mathbf{x}_{jm}^{(t)}$ given all other firms to be the same.

Now, denote the components of the likelihood increment that includes $\mathbf{x}$ to be

$$\begin{aligned}
& l_{\mathbf{x}}\left(\boldsymbol{\delta}_{jm}^{(t-1)}, \boldsymbol{\eta}_{jm}^{(t-1)}, \mathbf{w}_m, Q_m, \mathbf{X}_m^{(t-1)}; \boldsymbol{\theta}^{(t-1)}\right) \\
\equiv \quad & lnf_{(\boldsymbol{\xi},\boldsymbol{\upsilon})}\left(\left(\boldsymbol{\delta}_m^{(t-1)}, \boldsymbol{\eta}_m^{(t-1)}\right) - \left(\mathbf{X}_m^{(t-1)}\boldsymbol{\beta}^{(t-1)}, \mathbf{X}_m^{(t-1)}\boldsymbol{\gamma}^{(t-1)}\right)\right) + lnf_{\mathbf{x}}\left(\mathbf{X}_m^{(t-1)}|\mathbf{w}_m, Q_m, \boldsymbol{\theta}_{\mathbf{x}}^{(t-1)}\right)
\end{aligned}$$

I sample using the Random Walk Metropolis-Hastings algorithm.

Let $\mathbf{X}_m^{(t,j,k)}$ be the matrix with its jth column, $\mathbf{x}_{jm}^{(t,k)}$, being a vector whose lth element is:

$$\mathbf{x}_{jm,l}^{(t,k)} = \begin{cases} x_{jm,l}^{(t)} & \text{for } l < k \\ x_{jm,l}^{(t-1)} & \text{for } l \geq k \end{cases}$$

and let $\mathbf{X}_m^{(t,j,k)\dagger}$ be the vector whose lth element of jth column is:

$$\mathbf{x}_{jm,l}^{(t,k)\dagger} = \begin{cases} x_{jm,l}^{(t)} & \text{for } l < k \\ x_{jm,l}^{(t-1)} + \epsilon_N,\ \ \epsilon_N \sim N\left(0, \tau_{xk}\right) & \text{for } l = k \\ x_{m,l}^{(t-1)} & \text{for } l > k \end{cases}$$

Then, draw uniform distribution $\epsilon_U \sim U\left[0, 1\right]$, and let

$$\mathbf{x}_{jm,k}^{(t)} = \begin{cases} = \quad \mathbf{x}_{jm,k}^{(t,k)\dagger} & \text{if } \epsilon_U \leq \frac{exp\left(l_{\mathbf{x}}\left(\boldsymbol{\delta}_{jm}^{(t-1)}, \boldsymbol{\eta}_{jm}^{(t-1)}, \mathbf{w}_m, Q_m, \mathbf{X}_m^{(t,j,k)\dagger}; \boldsymbol{\theta}^{(t-1)}\right)\right)}{exp\left(l_{\mathbf{x}}\left(\boldsymbol{\delta}_{jm}^{(t-1)}, \boldsymbol{\eta}_{jm}^{(t-1)}, \mathbf{w}_m, Q_m, \mathbf{X}_m^{(t,j,k)}; \boldsymbol{\theta}^{(t-1)}\right)\right)} \\ = \quad \mathbf{x}_{jm,k}^{(t,k)} & \text{if otherwise} \end{cases}$$

2. Sampling δ_{jm} for $jm \notin \mathcal{E}$: conduct the following sampling procedure for $j = 1, \ldots, m$, for $jm \in \mathcal{E}$, Unlike section 4, p and s cannot be expressed in a closed form, and therefore I sample $\left(\delta_{jm}, \eta_{jm}\right)$ Given $\left(\mathbf{x}_m^{(t)}, \mathbf{w}_m, Q_m; \boldsymbol{\theta}^{(t-1)}\right)$, I augment the missing $\left(\delta_{jm}, \eta_{jm}\right)$ for each nonentrant firm $jm \notin \mathcal{E}$. I do so by using the Metropolis-Hastings algorithm. Let $\boldsymbol{\delta}_m^{(t,j)}$ be the vector whose lth element is:

$$\delta_{lm}^{(t,j)} = \begin{cases} \delta_{lm}^{(t)} & \text{for } l < j \\ \delta_{lm}^{(t-1)} & \text{for } l \geq j \end{cases}$$

and let $\boldsymbol{\delta}_m^{(t,j)\dagger}$ be the vector whose jth column is:

$$\delta_{lm}^{(t,j)\dagger} = \begin{cases} \delta_{lm}^{(t)} & \text{for } l < j \\ \delta_{lm}^{(t-1)} + \epsilon_N,\ \epsilon_N \sim N(0, \tau_{\delta k}) & \text{for } l = j \\ \delta_{lm}^{(t-1)} & \text{for } l > j \end{cases}$$

Let

$$\begin{aligned} & l_{(\delta,\eta)}\left(\boldsymbol{\delta}_m, \boldsymbol{\eta}_m, \mathbf{w}_m, Q_m, \mathbf{X}_m; \boldsymbol{\theta}\right) \\ & \equiv lnf\left(\mathbf{X}_m | \mathbf{w}_m, Q_m\right) + lnf_{(\boldsymbol{\xi},\boldsymbol{\upsilon})}\left(\boldsymbol{\delta}_m - \mathbf{X}_m\boldsymbol{\beta}, \boldsymbol{\eta}_m - \mathbf{X}_m\boldsymbol{\gamma}\right) + ln\left(J\left(\mathbf{e}_m\right)\right) \\ & + lnPr\left(\mathbf{e}_m | \mathbf{Z}_m\right) \end{aligned} \tag{59}$$

where

$$\mathbf{Z}_m \equiv \left(\mathbf{X}_m, \boldsymbol{\delta}_m, \boldsymbol{\eta}_m, w_m, r_m, Q_m\right).$$

Then, draw uniform distribution $\epsilon_a \sim U[0,1]$, and let

$$\boldsymbol{\delta}_m^{(t,j)} = \begin{cases} \boldsymbol{\delta}_m^{(t,j)\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l_{(\delta,\eta)}\left(\boldsymbol{\delta}_m^{(t,j)\dagger}, \boldsymbol{\eta}_m^{(t-1)}, \mathbf{w}_m, Q_m, \mathbf{X}_m^{(t)}; \boldsymbol{\theta}^{(t-1)}\right)\right)}{exp\left(l_{(\delta,\eta)}\left(\boldsymbol{\delta}_m^{(t,j)}, \boldsymbol{\eta}_m^{(t-1)}\ \mathbf{w}_m, Q_m, \mathbf{X}_m^{(t)}; \boldsymbol{\theta}^{(t-1)}\right)\right)} \\ \boldsymbol{\delta}_m^{(t,j)} & \text{if otherwise} \end{cases}$$

3. Sampling η_{jm} for $jm \notin \mathcal{E}$: conduct the following sampling procedure for $j = 1, \ldots, m$, for $jm \in \mathcal{E}$. Let $\boldsymbol{\eta}_m^{(t,j)}$ be the vector whose lth element is:

$$\eta_{lm}^{(t,j)} = \begin{cases} \eta_{lm}^{(t)} & \text{for } l < j \\ \eta_{lm}^{(t-1)} & \text{for } l \geq j \end{cases}$$

and let $\boldsymbol{\eta}_m^{(t,j)\dagger}$ be the vector whose jth column is:

$$\eta_{lm}^{(t,j)\dagger} = \begin{cases} \eta_{lm}^{(t)} & \text{for } l < j \\ \eta_{lm}^{(t-1)} + \epsilon_N,\ \epsilon_N \sim N(0, \tau_{\eta k}) & \text{for } l = j \\ \eta_{lm}^{(t-1)} & \text{for } l > j \end{cases}$$

Then, draw uniform distribution $\epsilon_a \sim U[0,1]$, and let

$$\boldsymbol{\eta}_m^{(t,j)} = \begin{cases} \boldsymbol{\eta}_m^{(t,j)\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l_{(\delta,\eta)}\left(\boldsymbol{\delta}_m^{(t)},\boldsymbol{\eta}_m^{(t,j)\dagger},\mathbf{w}_m,Q_m,\mathbf{X}_m^{(t)};\boldsymbol{\theta}^{(t-1)}\right)\right)}{exp\left(l_{(\delta,\eta)}\left(\boldsymbol{\delta}_m^{(t)},\boldsymbol{\eta}_m^{(t,j)}\,\mathbf{w}_m,Q_m,\mathbf{X}_m^{(t)};\boldsymbol{\theta}^{(t-1)}\right)\right)} \\ \boldsymbol{\eta}_m^{(t,j)} & \text{if otherwise} \end{cases}$$

Note the difference between the data augmentation in the monopoly case and here in duopoly case. In the monopoly case, I sampled (s_m, p_m) as the candidate draw for the Metropolis-Hastings candidate draw and then, obtained (δ_m, η_m) by inversion. Let us consider the case of $\mathbf{e}_m = (1,0)$ to explain the reason why I changed the algorithm. In this case, (δ_{1m}, η_{1m}) can be obtained by using inversion from the data on (s_{1m}, p_{1m}), and thus, it is more straightforward to draw (δ_{2m}, η_{2m}), and then, compute (s_{jm}, p_{jm}), $j = 1, 2$ under various monopoly/oligopoly situations to evaluate $Pr(\mathbf{e}_m|\mathbf{Z}_m)$.

5.3.2 Estimation of the structural parameters

The estimation of the structural parameters is identical to section 4.

5.4 Monte-Carlo Simulation Result

I conducted a Monte-Carlo simulation to assess the accuracy of the estimation method detailed in this section in recovering the true structural parameters. For this purpose, I created 2,000 hypothetical markets, each with two potential entrants. Each firm had a 50% chance of deciding whether to enter the market before their competitor. The remaining aspects of the Monte-Carlo simulation follow the same procedure as outlined in the analysis for Table 5.

The statistics of the posterior distributions are presented in Table 8, and the MCMC samplings are shown in figures 7 and 8.

The results show that the parameters were estimated with high accuracy, demonstrating the effectiveness of the proposed estimation procedure.

5.5 Extension to the BLP model

The argument so far used the logit demand model developed by Berry (1994). However, their model is known to have two major limitations. First, the model predicts an unrealistic substitution pattern. The price market share elasticity is only a function of prices and market shares, implying that all other variables, such as product and consumer characteristics, are irrelevant, which is clearly a strong restriction. Second, the Independence from Irrelevant Alternatives assumption is likely to be violated in the real world, which leads to unreliable counterfactual analysis.

The estimation method developed by Berry et al. (1995) is an extension of Berry (1994), which overcomes the two difficulties. The BLP model allows the consumer characteristics to enter the market share equation by making the structural parameters to have a distribution. One of the difficulties in implementing the BLP model is that now the market share function does not have a closed form expression and thus researchers needs to simulate the random coefficient variables to estimate the parameters.

Below, I show how to do the estimation when we adopt the BLP model to the demand side.

5.5.1 Economic Model[11]

The BLP model can be derived by adding consumer demographics to the logit demand model. Specifically, I consider the following utility function,

$$\begin{gathered} u_{ijm} = x_{jm}\beta_i - \alpha_i p_{jm} + \xi_j + \epsilon_{ijm}, \\ i = i, \ldots, I_m, j = 1 \ldots, J_m, m = 1, \ldots, M \end{gathered} \tag{60}$$

where

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i \tag{61}$$

D_i is a $d \times 1$ vector of demographic variables and Π is a matrix of parameters. The variable is assumed to have different distributions $F(D)$ in different markets. The distributions in each market can be identified non-parametrically. This variable represents the heterogeneity of consumers in each market, such as gender, age, education level, income, etc. The utility from the outside option is normalized to be ϵ_{i0m}.

Due to the form of the utility function, the market share function and the marginal revenue function does not have a closed formula.

$$s_{jm} = \int \frac{\exp(x_{jm}\beta_i - \alpha_i p_{jm} + \xi_{jm})}{1 + \sum_{r=1}^{J_m} \exp(x_{rm}\beta_i - \alpha_i p_{rm} + \xi_{rm})} \, dP(D) \tag{62}$$

$$MR_{jm} = p_{jm} + s_{jm} \left(\int \alpha_i s_{ijm}(1 - s_{ijm}) \, dP(D) \right)^{-1} \tag{63}$$

[11]This section uses notations similar to Nevo (2001).

Even though the marginal revenue function does not have a closed formula, Byrne et al. (2022) showed that Π is identified by equalizing MR and MC. Thus including Π in the utility function does not affect my identification strategy; (26) and (29) still identifies the price coefficient, labor and capital coefficients, and all random coefficients, and the rest of the parameters are identified using the logic of section 4.1.2.

5.5.2 Estimation

To derive the likelihood function, I make slight changes to (57):

$$
\begin{aligned}
l_m = \sum_{j}^{J_m} \left(-ln\pi - \frac{1}{2} ln\left|\mathbf{\Sigma_u}\right| - \frac{1}{2}\mathbf{u}'_{jm}\mathbf{\Sigma}_{\mathbf{u}}^{-1}\mathbf{u}_{jm} \right) \cdot I\left(mj \in \mathcal{E}\right) \\
+ lnf\left(\mathbf{X}_m | w_m, r_m, Q_m\right) + lnf_{(\boldsymbol{\xi},\boldsymbol{\upsilon})}\left(\boldsymbol{\delta}_m - \mathbf{X}_m\boldsymbol{\beta}, \boldsymbol{\eta}_m - \mathbf{X}_m\boldsymbol{\gamma}\right) + lnPr\left(\mathbf{e}_m | \mathbf{Z}_m\right)
\end{aligned}
\tag{64}
$$

where the component of $\mathbf{u}_{jm}$ is defined as

$$
\begin{aligned}
C_{jm} &= (\alpha_c + \beta_c) q_{jm} \left(p_{jm} + s_{jm} \left(\int \alpha_i s_{ijm}(1 - s_{ijm})\, d\hat{P}(D) \right)^{-1} \right) \\
C_{Ljm} &= \alpha_c q_{jm} \left(p_{jm} + s_{jm} \left(\int \alpha_i s_{ijm}(1 - s_{ijm})\, d\hat{P}(D) \right)^{-1} \right).
\end{aligned}
\tag{65}
$$

$\hat{P}$ shows the empirical distributions for D. The estimation of the structural parameters proceeds just like section 5.3, the only difference is that I replace the integrals of equation (65) with samplings from the empirical distributions.

5.5.3 Monte-Carlo Results

This section conducts a Monte-Carlo simulation that estimates the structural parameters in a BLP model. In the simulation, the utility function looks like the following:

$$
u_{ijm} = x_{jm}\beta - \alpha p_{jm} + \beta_x \pi_i x_{jm} + \alpha_p \pi_i p_{jm} + \xi_{jm} + \epsilon_{ijm} \tag{66}
$$

The variable π is sampled from a Bernoulli distribution, generating either 1 or 0 with equal probability, representing male and female, respectively. In the utility function, the terms $\beta_x \pi_i x_{jm}$ and $\alpha_p \pi_i p_{jm}$ capture the fact that males and females have different utility structures. The variable π is sampled 50 times for each market.

The demand-side parameters of interest are thus β, α, β_x, and α_p. The cost-side setup is identical to table 1. I run the MCMC 30,000 iterations and used the first 15,000 as burn-in period. I followed the estimation procedure discussed in the previous subsection and report the results in table 9. Each iterations' samplings are drawn in figure 9 and 10.

The results show that the estimation strategy identifies the structural paremeters.

5.6 BLP with micromoments

Petrin (2002)'s paper was the first to extend BLP's work by incorporating consumers' individual characteristics into the model using micromoment conditions. This innovation allowed Petrin to reduce standard errors and achieve more precise estimates of the structural parameters. Consequently, this method has gained popularity in demand analysis research, as evidenced by works such as Ching et al. (2015), Miller and Weinberg (2017), Hackmann (2019), Farronato and Fradkin (2022), and many others.

In this subsection, I conduct a Monte-Carlo simulation to demonstrate that it reduces the standard errors of random coefficient parameters. The utility function I consider is identical to the one in a BLP model. I add the following two moment conditions to align the predicted and observed number of male and female consumers:

$$
\begin{aligned}
E[consumer\ i\ is\ male | consumer\ i\ purchases\ product\ j] &= \#\ of\ male\ purchasing\ product\ j \\
E[consumer\ i\ is\ female | consumer\ i\ purchases\ product\ j] &= \#\ of\ female\ purchasing\ product\ j
\end{aligned}
\tag{67}
$$

To adapt the micromoments to the MCMC estimation approach, I convert the moment conditions to quasi-likelihood functions, as discussed in Tanaka (2020). Let $m(\theta)$ be the moment such that $E[m(\theta)] = 0$. Assuming there are J markets in total and only one potential entrant for simplicity, I define the quasi-likelihood function for each moment condition as follows:

$$
q(\theta|D) = \left(\frac{2\pi}{J}\right)^{-J/2} exp\left[-\frac{1}{2}\bar{m}(\theta)'\bar{m}(\theta))\right] \tag{68}
$$

where $\bar{m}(\theta) = \frac{1}{J}\sum^{J} m(\theta)$, is the sample analog of the moment condition, and D is the data. I add equation (68) to likelihood function (64) and estimate the parameters.

Table 10 shows the estimation results, and Figures 11 and 12 illustrate the MCMC process. The results indicate that the moments help minimize the variance of the posterior distributions of the random coefficient parameters.

6 Empirical Analysis of the Skilled Nursing Facility Market

In this section, I apply the analytical method developed in the previous sections to estimate the structural parameters of the skilled nursing facility (SNF) market and analyze the effect of the Certificate of Need(CON) law.

With the increasing elderly population, Wisconsin has seen its state expenditures for nursing homes increase significantly. To curb these escalating costs, the state implemented the CON law.[12] This legislation restricts the construction of new nursing facilities and the expansion of existing ones. Although the CON law may have reduced government spending, a large body of literature has studied its negative effects due to excess demand in the industry.

First, Gruenberg and Willemain (1982), Gertler (1989), and Gertler (1992) use both economic theory and empirical studies to verify the relationship between excess demand and patients' access to nursing homes. Limited supply led Medicaid patients to be rationed out, since their reimbursement rate was significantly lower than that of private-pay patients (PPPs). This is problematic for at least a couple of reasons. Firstly, low-income potential patients who need professional assistance may not be able to get any treatment. Secondly, as Ettner (1993) pointed out, the average hospital cost per patient-day for a semiprivate hospital room was $465 in 1985, while the average nursing homes' private patient price was only $1,456 per month. Since 40% of nursing home residents come directly from hospitals, they note that supply constraints may create inefficient government spending. Kotschy and Bloom (2022) examined data from 30 developed countries and found that difficulty in accessing nursing homes is a common issue globally. Second, Nyman (1985), Nyman (1988a), Nyman (1988b), Harrington et al. (2000), Horn et al. (2005), Lin (2015), Lu et al. (2021), and Kunz et al. (2024) consider various measures of nursing home quality and discuss its relationship with excess demand, most of them reporting a negative relationship between the two variables. Nyman (1988a) points out that nursing homes' low quality was a well-known issue, as officially reported in the senate report in 1994 Due to low Medicaid reimbursement and flat rates, reducing costs rather than engaging in quality competition provided greater benefits for nursing homes. Gupta et al. (2024) discuss the relationship between health condition and private equity and finds that private equity ownership increases mortality rate

[12] The rationale of the CON law and its history is described in the following FTC document: https://www.ftc.gov/sites/default/files/documents/reports/improving-health-care-dose-competition-report-federal-trade-commission-and-department-justice/040723healthcarerpt.pdf

by 10%. Third, Norton (1992), Cohen and Spector (1996) , Grabowski (2001), and Grabowski and Angelelli (2004), focus on Medicaid reimbursement rates and their implications in relation to market outcomes. Especially, Grabowski (2001) and Grabowski and Angelelli (2004) find a positive effect of an increase in Medicaid reimbursement rate on health outcomes, although the CON may mitigate this effect. Fourth, Nyman (1994) explores the relationship between price and excess demand, which they find a positive relationship between the two variables. Bardey and Siciliani (2021) uses two-sided economic model and finds that the profits and wages for nurses become lower when the prices are regulated, and Heger et al. (2022) uses exogenous variation in Swiss care price regulation and finds that higher prices leads to higher staffing ratio. Yang et al. (2022) offer a literature review regarding pricing behavior and concentration in the industry.

However, as pointed out in Ching et al. (2015), most of the empirical strategies in the existing literature use reduced-form estimation methods, which have several limitations. First, they cannot estimate social welfare. Second, their estimation strategies often suffer from endogeneity issues. For example, Gertler (1989) analyzes the effect of excess demand on quality by running regressions that include measures of excess demand and the Herfindahl-Hirschman Index (HHI) in the right hand side to control for market structure. However, as Miller et al. (2022) points out, both excess demand and HHI are market outcomes influenced by observed and unobserved shocks. As a result, it is impossible to establish causal relationship between the two endogenous variables. Analysis of effect of excess demand on market structure requires a structural approach.

Recent papers have begun using structural approaches to analyze the effect of excess demand and the market structure. Ching et al. (2015) develop a static oligopoly model based on Berry et al. (1995) to estimate private-patient demand. They use the estimated parameters to quantify the rationing effect of the CON law. However, since they only model the demand side, they cannot estimate the producer-side welfare, which is also essential in discussing the CON law's effect. Moreover, their static model cannot fully analyse how profit competition between firms promotes quality competition. Hackmann (2019) extends Ching et al. (2015)'s study by endogenizing nursing homes' quality and explicitly modeling the cost side. In addition to the static models shown above, there are several papers that use dynamic models to study the market structure. Gowrisankaran and Town (1997) was one of the pioneers that used a full dynamic model that levaraged Ericson and Pakes (1995) model to estimate demand, cost, entry/exit cost, income elasticity, and other parameters of the hospital market. This paper shows groundbreaking results and implications that has a lot of common things with the nursing home industry, but it is not possible to discuss the level of rationing with this economic model. Lin (2015) establishes a dynamic model following Ericson and Pakes (1995), where nursing homes make dynamic decisions about their treatment quality. They use panel data to identify the competition effect of profit and entry cost. Nonetheless, their model is not suitable for analyzing the CON law as they do not model consumer choices and the cost of production, making it impossible to quantify the social effect of the CON law or the rationing effect. Grant et al. (2022) uses Germany nursing homes' market data and develops an entry model to study the competition of for-profit and non-profit nursing homes. Since they simplify the profit function in order to focus on dynamic entry game behavior, and Germany do not have laws similar to the CON law, measuring the effect of the CON using this model is not suitable. Overall, there are no research papers that thoroughly estimate the CON law's effect and quality competition effect incorporating entry-exit models.

The research most closely related to mine is Hackmann (2019), who developed a model similar to that of Ching et al. (2015). In their study, they endogenize nursing home quality, i.e. the number of nurses per patient, and use a constant marginal cost function, estimating the model using instrumental variables (IV). However, their model has several limitations. First, their identification strategy using IV may be biased. They use Medicaid patients' reimbursement rate as an instrument for the demand shock, assuming that the cost variation is orthogonal to unobserved preference shocks in the given nursing home county. However, this assumption does not always hold, especially if the cost shock is serially correlated or correlated with the cost structure, which cannot be verified. Secondly, they do not model the entry behavior of nursing homes. Although the CON law restricts the construction of new nursing homes, the data in Table 1 show that the number of facilities has fluctuated significantly each year. Therefore, as discussed in Table 3, the structural parameters are biased when the entry behavior of firms is not taken into account. Third, they use a constant marginal cost function, which lacks flexibility due to the lack of cost data. As demonstrated in my research, nursing homes exhibit increasing returns to scale, which could bias their counterfactual simulations analyzing the impact of an increase in Medicaid patients' reimbursement rates on facility quality. Unlike Hackmann (2019), the model I use can estimate the cost function flexibly, as the estimation strategy does not require IV.

This paper uses the structural approach developed in the previous sections to answer the following two major questions prevalent in Wisconsin's skilled nursing home facility market:

1. What is the quantitative effect of the CON law on social welfare and rationing?

2. Does competition improve the quality of treatment?

Most of the previous studies mentioned so far have analyzed various types of nursing facilities including skilled nursing facilities, nursing homes connected to hospitals, and home health agencies without making clear distinctions between them. In my research, however, instead of analyzing all types of nursing homes, I focus only on skilled nursing facility due to data limitation. According to California Department of Aging,a typical resident of SNF is 'a person who is chronically ill or recuperating from an illness or surgery and needs regular nursing care and other health related services.' Unlike other types of nursing homes, a SNF 'is a temporary residence for patients undergoing medically necessary rehabilitation treatment.'

Following Ching et al. (2015), the model I use assumes that private-pay patients (PPP) do not face rationing, while Medicaid patients (MP) do face rationing. I use PPP's discrete choice outcomes to estimate the structural parameters and conduct counterfactual simulations of MP using the estimated structure.

The model assumes that there is only one market, the state of Wisconsin, which includes a total of 286 entrant SNFs and 286 unobservable nonentrant SNFs. However, since SNF's orders of entry are not observed, I need to assign probabilities of entry orders to firms as discussed in section 5.1 to estimate the structural parameters, but calculating all possible orders of entry for the existing 572 SNFs is infeasible, as it would require computing 572! (factorial) possible orders. Instead, I assume that SNFs make entry-exit decisions in a monopolistic competition environment. In the economic model used in this analysis, in period 0, SNFs calculate their profit according to an unknown strategy and those who can achieve a positive profit announce that they will enter the market and their price. In period 1, SNFs make entry and exit decisions again, using the announced information to calculate their profits. During this period, SNFs face monopolistic competition, where they assume that their entry-exit decisions do not affect other SNF's entry-exit decisions and their prices. Consequently, each SNF makes its entry and exit decisions based on the assumption that the actions and prices of its competitors remain the same as in period 0. This simplification makes the estimation tractable because each firms' entry-exit decision is no longer a function of other SNF's entry-exit decisions. The monopolistic competition assumption is justified by the fact that each firm's market share is very small, given the presence of 572 SNFs in the market. The outcome of period 1 is observed by the econometrician in period 2. I estimate the model using Gibbs-in-Metropolis-Hastings MCMC.

This article also contributes to the literature on the estimation of nursing home cost functions. The literature has used a variety of methods to measure the cost structure of nursing homes. Vitaliano and Toren (1994), Hofler and Rungeling (1994), and Mutter et al. (2013) use the stochastic frontier approach to examine the inefficiency of nursing homes. Bekele and

Holtmann (1987), Gertler and Waldman (1992), Filippini (2001), and Giorgio et al. (2016) use the translog cost function structure to estimate flexible cost functions. Dudzinski et al. (1998) use the Hedonic approach to find nursing home's returns to scale, and Christensen (2004) use quantile regression with a translog form to discuss the heterogeneity in efficiency. Some other papers use reduced-form approaches; Knox et al. (2004) use regression models to discuss the heterogeneity of efficiencies across different types of nursing homes, and Murmann et al. (2023) use the propensity-score method to discuss the cost-effectiveness of transitional care in Canada.

However, to the best of the author's knowledge, none of them address the endogeneity of quantity. Production quantity is determined by a firm's profit maximization problem. Therefore, firms with higher marginal costs (or higher cost shocks) produce fewer products. Consequently, there is likely a correlation between observed production quantity and unobserved heterogeneity in cost structures or cost shocks. This issue has been overlooked in the literature due to a lack of valid instrumental variables. In contrast, the estimation method used in this article circumvents this problem by canceling out the unobserved cost components from the estimation equations. I estimate the cost function derived from the Cobb-Douglas production function and discuss its implications.

This paper is also one of the few to provide estimates of entry costs. Most papers rely on the dynamic behaviour of firms to identify them using panel data. They find different estimation results regarding entry costs; Gowrisankaran and Town (1997) finds the entry cost is statistically insignificant to zero, while Lin (2015) and Grant et al. (2022) finds a large entry cost. My article identifies the entry cost in a static game from a novel perspective; it identifies the entry cost from observed firms' profit distribution. Firms having large profit implies high entry cost in my model because otherwise it cannot explain other potential entrants not entering the market despite the market being profitable. I show that the entry cost is relatively small compared to the total cost, as the mean of the profit distribution of SNF is close to zero.

The estimation results of my model carry significant policy implications. Firstly, the labor and capital elasticities are 0.698 and 0.662, respectively, summing to 1.361. This result implies increasing returns to scale, a finding corroborated by previous research(e.g. Bekele and Holtmann (1987), Filippini (2001)). Such insights are crucial for policymaking, indicating that government initiatives to invest in SNFs would be more cost-effective if focused on expanding larger establishments rather than building many smaller ones. Secondly, the estimated daily entry cost for a nursing home is $10.9, amounting to an annual cost of $4,102. This figure is considerably lower compared to the total operational costs, which average $11,704 daily, or approximately $4.27

million annually. This disparity likely stems from intense competition within the SNF sector, with the entry cost in my model identified from firms' profitability.

In conducting a series of counterfactual simulations using estimated demand, cost, and entry parameters, I first evaluated the impact of the Certificate of Need (CON) law on rationing effects and social welfare. Assuming the absence of the CON law, an additional 17,287 patients would have access to SNFs. Furthermore, both consumer and producer surplus would increase by $868 million and $165 million annually, respectively. Meanwhile, government expenditure would rise by $700 million, leading to an annual increase in social welfare of $333 million. This result supports the findings of Ching et al. (2015), which suggest that current policy results in a net welfare loss. Additionally, I analyzed the distribution of quality among nursing homes, comparing current entrants to potential entrants. Utilizing Ching et al. (2015)'s measure of nursing home quality, defined as $\mathbf{x}\boldsymbol{\beta}+\xi$, I discovered that the average quality measure of entrants is 10% higher than that of potential entrants. This indicates that entry-exit competition significantly improves quality, offering a potential solution to the problem of substandard quality in nursing homes.

In the upcoming subsections, I first offer a detailed overview of the nursing home market structure and regulatory environment in Wisconsin. Subsection 6.2 delves into the dataset I used and present some key descriptive statistics that provide insight into the market. Following that, Subsection 6.3 outlines the empirical strategy I employed to analyze the data and the specific methodologies behind my analysis. Finally, subsection 6.4 discusses the results of my estimations, highlighting the significant findings and their implications.

6.1 Structure of Wisconsin's Nursing Home Market

This section closely follows Ching et al. (2015).

6.1.1 Certificate of Need

Wisconsin has installed the CON law for nursing homes since 1980 under a clear purpose to manage the state budget, as is written in the state statutes: 'it exists in order to enable the state to budget accurately for medical assistance and to allocate fiscal resources most appropriately...'[13]. Wisconsin has the statewide bed limit of 51,795 in 1999. Wisconsin also limits the number of beds in each county; they allow nursing homes to increase their total number of beds when other nursing home closes. As a result, the occupancy tend to be high, with an average of 91%.

[13]Wisconsin Statutes Chapter 150

Moreover, when somebody wants to build a nursing facility or a facility wants to expand its capacity, they need to submit an application along with an application fee which is equal to 0.37% of the estimated project cost, but no less than $1,850 and no more than $37,000[14].

Due to this restriction, none of the nursing homes present from 1998 to 2002 changed their bed capacity after entering the market. However, the number of nursing homes did change during that period. Table 1 illustrates trends in the number of nursing homes. Although the Certificate of Need (CON) regulation restricts changes in capacity and new entries for nursing homes, I observed fluctuations in their numbers for the latter condition only. This observation supports my economic model, which allows the entry and exit of firms in the market.

6.1.2 Quality of Care

Wisconsin has minimum staffing requirements for the number of nurse hours per bed[15]. Specifically for every nursing facility,

- For each resident in need of intensive skilled nursing care, 3.25 hours per day, of which a minimum of 0.65 hour shall be provided by a registered nurse or licensed practical nurse.
- For each resident in need of skilled nursing care, 2.5 hours per day, of which a minimum of 0.5 hour shall be provided by a registered nurse or licensed practical nurse.
- For each resident in need of intermediate or limited nursing care, 2.0 hours per day, of which a minimum of 0.4 hour shall be provided by a registered nurse or licensed practical nurse.

Furthermore, the Wisconsin Administration Code Chapter HFS 132 mandates that nursing homes provide the same quality of care to all patients, regardless of the payment source or payment amount. Empirical research by Grabowski et al. (2008) confirms that patients with different payment sources indeed receive the same quality of care.

6.1.3 Patients

In this research, I focus on two primary types of patients due to their differing payment methods: private-pay patients and Medicaid patients. PPPs cover their expenses out-of-pocket, while MPs

[14] Wisconsin Statute, https://ij.org/report/conning-the-competition/state-profile/wisconsin/

[15] Wisconsin Statute, https://docs.legis.wisconsin.gov/statutes/statutes/50

receive government reimbursement for their care, with no out-of-pocket costs involved. Medicaid eligibility requires meeting strict income and asset qualifications, suggesting that switching between these patient types is not feasible.

An additional assumption in my analysis is that PPPs do not experience rationing. Data indicates that PPPs pay a significantly higher price for care compared to the reimbursement rate for MPs. Consequently, I posit that SNFs prioritize filling their beds with PPPs first. Once all PPPs have selected their accommodations, SNFs then allocate remaining beds to MPs. This preference for PPPs is supported by findings from Ettner (1993), who noted longer waiting times for MPs based on waitlist data. Furthermore, according to my data from Wisconsin in 2000, 22% of patients in SNFs were PPPs, and 69% were MPs. These findings suggest that its impractical for SNFs to rely solely on PPPs to fill their beds, and no SNF in the dataset had all beds occupied exclusively by PPPs.

In summary, my model assumes that only MPs face rationing due to the preference of SNFs for PPPs. This leads to the strategy where SNFs may equate marginal revenue (MR) to marginal cost (MC) to attract PPPs and maximize their profit, but this balance might not hold when accommodating MPs. Therefore, my analysis focuses on the choices and associated costs of PPPs to estimate the model parameters.

6.2 Data

I combine four data sources: 1999 Wisconsin Annual Survey of Nursing Homes, Skilled Nursing Facility Cost Report, 1999 Wisconsin Health Survey, and 2000 Census of Population.

The 1999 Wisconsin Annual Survey of Nursing Homes provide aggregate data of consumer choices and characteristics of nursing homes. It contains the number of PPP, MP, capacity of each nursing home, per-diem rate, and characteristics such as total number of nurses per bed, licensed practical nurses, nursing assistants, therapist, and so on.

The Skilled Nursing Facility Cost Report provides detailed information on each nursing SNF's overall expenses, including total labor and capital costs. Additionally, this report, along with the 1999 Wisconsin Annual Survey of Nursing Homes, records the number of nursing staff employed. However, there are instances where the figures between the two sources does not match. In such situations, I refer to the numbers provided by the 1999 Wisconsin Annual Survey of Nursing Homes.[16]

[16]Even though the demand and cost data include information about three types of nursing facilities' (skilled nursing facilities (SNFs), nursing homes built along hospitals, and home health agencies), I have chosen to focus solely on SNFs for this analysis. This is because nursing homes built along hospitals and home health agencies

I use the remaining two datasets to derive the total potential patients in Wisconsin. 2000 Census of Population reveals the number of elderly people conditional on sex and income level relative to the poverty line. 1999 Wisconsin Health Survey shows the percentage of elderly people whose health condition is poor/fair/good/excellent.

Some nursing homes data are not reported in 1999 Wisconsin Annual Survey or Skilled Nursing Facility Cost. I use the data of nursing homes that reports their data on both data sources. Moreover, I exclude SNFs with zero MP, as they primarily treat Medicare patients or patients with mental disorders. As a result, there are 286 SNFs used for the analysis.

The descriptive statics are given in table 11. The variables listed in the table are observable for entrants only, except for the rental rate, for which I used the national policy rate in December 1999. I assume that the number of nonentrants is equal to the number of entrants. The mean of the observed hourly wage rate is used as the hourly wage rate for non-entrants. The total number of potential MPs is the number of people living in Wisconsin who are >65 years old and whose income is below twice the Wisconsin poverty line with unhealty condition, resulting in a total of 52,028 potential patients.

There are two types of product characteristics x. The first are variables that take only positive values, such as nurses per bed, which I assume that they follow the log-normal distribution. The second type of variables are dummies that indicate whether the nursing homes are government-owned facilities, organized by nonprofit agencies, or organized by for-profit agencies. All nursing homes fall into one of these three categories. I assume they follow the categorical distribution. Going forward, I will use the terms 'nursing homes', 'SNFs', and 'firms' interchangeably in this article.

6.3 Economic Model

This section discusses the economic model of the SNF market. I assume that the entire state of Wisconsin is the market, and that potential patients in the state choose whether or not to enter a nursing home in the state.

There are 286 entrant nursing homes and 286 non-entrants in the Wisconsin market. Therefore, computing the entry equilibrium by assigning probabilities to entry orders is infeasible because it involves computing 572! patterns of sequential entry decisions and 2^{572} patterns of hypothetical nursing home profits. Therefore, I make a simplifying assumption: each nursing

report their total costs for all medical services combined, making it difficult to determine the specific costs associated with nursing home services.

home faces monopolistic competition[17]. Specifically, I assume that firms in this empirical model make decisions in the following time frame:

- Period 0: Firms follow an unknown strategy and calculate their profit. Those who can achieve a positive profit announce that they will enter the market and their price.
- Period 1: Firms make entry/exit and pricing decisions simultaneously again, assuming that entry and pricing decisions in Period 0 remain fixed regardless of the firm's decisions. As a result of this assumption, I do not need to model the order of entry of firms, which makes the estimation tractable. The econometrician needs only to compute the best response price and its associated profit, given all other firms' characteristics and decisions fixed, which can be computed by solving a one-dimensional profit minimization problem.
- Period 2: Firms' entry/exit and pricing equilibria are realized and the econometrician observes the equilibrium.

I use this simplified competition model to construct the econometric model described in the following subsections.

6.3.1 Demand, Supply, and Entry

The utility functions of PPP and MP are given as follows:

$$\begin{aligned} u_{ij}^p &= -\alpha p_j + \mathbf{x}_j\boldsymbol{\beta} + \xi_j + \epsilon_{ij} \\ u_{ij}^m &= \kappa(\mathbf{x}_j\boldsymbol{\beta} + \xi_j) + \epsilon_{ij} \end{aligned} \tag{69}$$

where the definition of the variables follows the previous chapters. κ shows the difference of the utility perceived between PPP and MP. MP does not have to pay the fee by themselves and thus the price coefficient is not included in their utility function.

The last term follows an i.i.d. extreme value distribution, and thus the market share of firm j is determined by the logit formula for PPP:

$$s_j^p = \frac{exp\left(-p_j\alpha + \mathbf{x}_j\boldsymbol{\beta} + \xi_j\right)}{1 + \sum_k exp\left(-p_k\alpha + \mathbf{x}_k\boldsymbol{\beta} + \xi_k\right)} \tag{70}$$

[17] The concept of monopolistic competition is not a new one. For instance, Melitz (2003) develops a trade model in which firms assume the price levels of their competitors are fixed.

Since MP could face rationing, their market share cannot be expressed as a simple logit model like PPP's market share function. To deal with this issue, first I divide the Medicaid population into R groups, $\{M_1^m, M_2^m, \ldots, M_R^m\}$. M_1^m show the Medicaid population who can choose any nursing home, M_2^m show the population who can choose any nursing homes except for the one which has been taken away from M_1^m, and so on. Therefore the aggregate demand for MP is calculated as

$$n_j^m = \sum_r M_r^m \frac{\exp\left(\kappa(\mathbf{x}_j\boldsymbol{\beta} + \xi_j)\right)}{1 + \sum_{k \in J_r} \exp\left(\kappa(\mathbf{x}_k\boldsymbol{\beta} + \xi_k)\right)} \tag{71}$$

and the market share of firm j's MP patients is given as

$$s_{j,r}^m = \frac{\exp\left(\kappa(\mathbf{x}_j\boldsymbol{\beta} + \xi_j)\right)}{1 + \sum_{k \in J_r} \exp\left(\kappa(\mathbf{x}_k\boldsymbol{\beta} + \xi_k)\right)}. \tag{72}$$

The cost functions for PPP and MP are defined as:

$$\begin{aligned} C_j^{p*} &= \left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)} \times exp\left(\mathbf{x}_j\boldsymbol{\gamma} + v_j\right) q_{pj}^{(1/(\alpha_c+\beta_c))} \\ C_j^{m*} &= \left(\frac{w}{\alpha_c}\right)^{\alpha_c/(\alpha_c+\beta_c)} \left(\frac{r}{\beta_c}\right)^{\beta_c/(\alpha_c+\beta_c)} \times exp\left(\mathbf{x}_j\boldsymbol{\gamma} + v_j\right) q_{mj}^{(1/(\alpha_c+\beta_c))} \end{aligned} \tag{73}$$

where q_{pj} and q_{mj} shows the number of PPP and MP in nursing home j, respectively. However, the dataset only provides the combined total, $C_j^{p*} + C_j^{m*}$, without reporting the individual costs for PPP and MP separately. Thus, as a measure of the total cost for PPP, I use the total cost of the nursing home multiplied by each nursing home's share of PPP.

I also assume that the number of entrants and non-entrants is the same. Firms that announced that they enter the market in period 1, $j \in J_e$, decides to enter the market in period 2 if and only if its profit from PPP is positive:

$$\pi_j^p = p_j * MS^p * \frac{exp\left(-p_j\alpha + \mathbf{x}_j\boldsymbol{\beta} + \xi_j\right)}{1 + \sum_k exp\left(-p_k\alpha + \mathbf{x}_k\boldsymbol{\beta} + \xi_k\right)} - C^{p*}(q_{pj}) > EC - \epsilon_j + \epsilon_j \tag{74}$$

whereas the firms l who did not announce the entry decides to enter to the market in period 2 if and only if:

$$\pi_l^p = p_l * MS^p * \frac{exp\left(-p_l\alpha + \mathbf{x}_l\boldsymbol{\beta} + \xi_l\right)}{1 + \sum_k exp\left(-p_k\alpha + \mathbf{x}_k\boldsymbol{\beta} + \xi_k\right) + exp\left(-p_l\alpha + \mathbf{x}_l\boldsymbol{\beta} + \xi_l\right)} - C^{p*}(q_{pj}) > EC - \epsilon_l + \epsilon_l \tag{75}$$

6.3.2 Econometric Model

Following the previous sections, the regression equations to identify $(\alpha, \alpha_c, \beta_c)$ are:

$$
\begin{aligned}
C_j^p =& \left(\alpha_c + \beta_c\right) q_{pj} \left[p_j^p - \frac{1}{\left(1 - s_j^p\right)\alpha} \right] + u_j \\
C_{Lj}^p =& \alpha_c q_{pj} \left[p_j^p - \frac{1}{\left(1 - s_j^p\right)\alpha} \right] + u_{Lj}
\end{aligned} \tag{76}
$$

First, I estimate $(\alpha, \alpha_c, \beta_c)$ by applying ordinary least squares (OLS) to equation (76). Then, I estimate the remaining parameters, except κ, by constructing the likelihood function and performing Markov Chain Monte Carlo (MCMC) simulations using the following likelihood function[18]:

$$
\begin{aligned}
lnf\left(\mathbf{x}_j, \delta_j, \eta_j | \mathbf{w}, Q\right) =& lnf_{\mathbf{x}}\left(\mathbf{x}_j | \mathbf{w}_j, Q_j\right) + lnf_{(\xi,\upsilon)}\left(\delta_j - \mathbf{x}_j\boldsymbol{\beta}, \eta_j - \mathbf{x}_j\boldsymbol{\gamma}\right) + lnPr(Entry_j) \\
&+ \left[-\frac{K}{2} ln\pi - \frac{1}{2} ln\left|\boldsymbol{\Sigma}_c\right| - \frac{1}{2}\mathbf{u}_j'\boldsymbol{\Sigma}_c^{-1}\mathbf{u}_j \right]
\end{aligned} \tag{77}
$$

$$
\begin{aligned}
& l_j\left(s_j, p_j, w_j, r_j, Q_j, C_j, C_{Lj}, \mathcal{E}\right) \\
\equiv\ & l_{uj} + l_{hj} + l_{ej} + l_{\mathbf{x}j} \\
=\ & \left[-\frac{K}{2} ln\pi - \frac{1}{2} ln\left|\boldsymbol{\Sigma}_{\mathbf{u}}\right| - \frac{1}{2}\mathbf{u}_j'\boldsymbol{\Sigma}_{\mathbf{u}}^{-1}\mathbf{u}_j \right] I\left(j \in \mathcal{E}\right) \\
& + lnf_{(\xi,\upsilon)}\left(h\left(\left(s_j, p_j\right) | w_j, r_j, Q_j\right) - \left(\mathbf{x}_j\boldsymbol{\beta}, \mathbf{x}_j\boldsymbol{\gamma}\right)\right) + ln\left(J_{\left(\left(\xi_j, \upsilon_j\right) \to \left(s_j, p_j\right)\right)}\right) \\
& + ln\left(\frac{p_j q_j - C_j^* - E_C}{1 + exp\left(p_j q_j - C_j^* - E_C\right)} \right) I\left(j \in \mathcal{E}\right) + ln\left(\frac{1}{1 + exp\left(p_j q_j - C_j^* - E_C\right)} \right) I\left(j \notin \mathcal{E}\right) \\
& + lnf_{\mathbf{x}}\left(\mathbf{x}_j | w_j, r_j, Q_j, \boldsymbol{\theta}_{\mathbf{x}}\right)
\end{aligned} \tag{78}
$$

where the entry probability is given by (74) and (75).

I describe each component of the likelihood function below:

[18] These parameters are estimated first because they can be easily obtained using an OLS model, allowing me to explore different model setups while maintaining reasonable estimates of the three parameters. Estimating these parameters first and then the remaining parameters requires a two-step bias correction, as discussed in Duncan (1987), which is currently in progress.

First Component: The third component comes from equation $MR = MC$:

$$l_{cj} = -\frac{K}{2} ln\pi - \frac{1}{2} ln \left|\mathbf{\Sigma}_c\right| - \frac{1}{2}\mathbf{u}_m' \mathbf{\Sigma}_c^{-1} \mathbf{u}_m \tag{79}$$

Second Component: The second component, $lnf_{(\xi,\upsilon)}\left(h\left((s_j,p_j)\left|w_j,r_j,Q_j\right.\right)-(\mathbf{x}_j\boldsymbol{\beta},\mathbf{x}_j\boldsymbol{\gamma})\right)+$

$ln\left(J_{((\xi_j,\upsilon_j)\to(s_j,p_j))}\right)$ is the distribution of (ξ, υ), which takes the functional form similar to (40) and (41).

Third Component: The third component, $ln\left(\frac{p_jq_j-C_j^*-E_C}{1+exp\left(p_jq_j-C_j^*-E_C\right)}\right)I\left(j\in\mathcal{E}\right)+$ $ln\left(\frac{1}{1+exp\left(p_jq_j-C_j^*-E_C\right)}\right)I\left(j\notin\mathcal{E}\right)$ comes from the entry probability. The entry probability has a closed formula because the entry shock follows the type 1 i.i.d. extreme distribution and the monopolistic competition assumption.

Fourth Component: The last component is the likelihood of $(\mathbf{x}, \delta, \eta)$.

First, I define the distribution of $\mathbf{x}$. In this empirical analysis, I assume that continuous variables k of firm j follows the log-normal distribution:

$$x_{kj} \sim LN(\mu_{kj}, \sigma_{kj}) \tag{80}$$

while the dummy variables that indicate whether the nursing homes are governmental, non-profit, or for-profit follows the categorical distribution:[19]

$$(x_{kj}, x_{k+1j}, x_{k+2j}) \sim Categorical(p_{1k}, p_{2k}, p_{3k}). \tag{81}$$

ξ and υ follow the normal distribution.

After estimating the demand parameters in PPP and cost parameters, I estimate κ by equalizing the observed number of MP and predicted number of MP:

$$\hat{\kappa} = arg\min_{\kappa}\sum_j\left(n_j^m - \sum_r M_r^m \frac{\exp\left(\kappa(\mathbf{x}_j\hat{\boldsymbol{\beta}}+\xi_j)\right)}{1+\sum_{k\in J_r}\exp\left(\kappa(\mathbf{x}_k\hat{\boldsymbol{\beta}}+\xi_k)\right)}\right)^2 \tag{82}$$

where $\hat{\boldsymbol{\beta}}$ is the estimated $\boldsymbol{\beta}$.

[19] The parameters of those distributions are given as follows:
$\mu_{xk} \sim N(\overline{log(x_k)}, \sigma_{xk}^2/m)$,
$\sigma_{xk} \sim IVG(m/2, \sum_j(log(x_{jk}) - \mu_{kj})$,
$(p_{1k}, p_{2k}, p_{3k}) \sim Dirichlet(\sum x_{kj}, \sum x_{k+1j}, \sum x_{k+2j}))$

6.3.3 Estimation: Data Augmentation and MCMC

Data Augmentation Method

I start this subsection by explaining how to augment x based on their distribution.

There are two types of product characteristics as discussed in the previous subsection. To augment continuous x, first draw the proposal value $\mathbf{x}^{\dagger}$ as follows:

$$\begin{aligned} \mathbf{x}_{j-k}^{\dagger} &= \mathbf{x}_{jk} && (83) \\ x_{jk}^{\dagger} &= x_{jk} + \epsilon_N, \ \ \epsilon_N \sim N\left(0, \kappa_{xk}\right). && (84) \end{aligned}$$

where k is one of the variables in $\mathbf{x}_j$. While the proposal distribution becomes as follows for the dummy variables:

$$\begin{aligned} \mathbf{x}_{j-k}^{\dagger} &= \mathbf{x}_{jk} && (85) \\ x_{k}^{\dagger} &\sim Categorical\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) && (86) \end{aligned}$$

Then, draw uniform distribution $\epsilon_U \sim U\left[0,1\right]$, and let

$$\mathbf{x}_j = \begin{cases} = \ \mathbf{x}_j^{\dagger} & \text{if } \epsilon_a \leq \frac{exp\left(l\left(s_j, p_j, w_j, r_j, Q_j, \mathbf{x}_j^{\dagger}; \boldsymbol{\theta}\right)\right)}{exp\left(l\left(s_j, p_j, w_j, r_j, Q_j, \mathbf{x}_j; \boldsymbol{\theta}\right)\right)} \\ = \ \mathbf{x}_j & \text{if otherwise} \end{cases}$$

Now, to augment (s, p), generate the proposal values as

$$s_j^{\dagger} = s_j + \epsilon_{sj}, p_j^{\dagger} = p_j + \epsilon_{pj},$$

where $\epsilon_{sj} \sim N\left(0, \tau_s\right)$ and $\epsilon_{\tau j} \sim N\left(0, \tau_p\right)$.

Then, draw uniform distribution $\epsilon_a \sim U\left[0,1\right]$, and let

$$(s_j, p_j) = \begin{cases} = \ \left(s_j^{\dagger}, p_j^{\dagger}\right) & \text{if } \epsilon_a \leq \frac{exp\left(l\left(s_j^{\dagger}, p_j^{\dagger}, w_j, r_j, Q_j, \mathbf{x}_j; \boldsymbol{\theta}\right)\right)}{exp\left(l\left(s_j, p_j, w_j, r_j, Q_j, \mathbf{x}_j; \boldsymbol{\theta}\right)\right)} \\ = \ (s_j, p_j) & \text{if otherwise} \end{cases}$$

Estimation of Structural Parameters

Now I discuss how to estimate the structural parameters.

Let $\boldsymbol{\delta}$ be the vector whose j^{th} element is δ_j. Similarly, let $\mathbf{X}$, be the matrix whose j^{th} row is

$\mathbf{x}_j$. Then, draw $\boldsymbol{\beta}$ from $\boldsymbol{\beta} \sim N\left(\widehat{\boldsymbol{\beta}}, \sigma^2_{\boldsymbol{\beta}} \left(\mathbf{X}'\mathbf{X}\right)^{-1}\right)$, where

$$\widehat{\boldsymbol{\beta}} = \left[\mathbf{X}'\mathbf{X}\right]^{-1} \mathbf{X}'\boldsymbol{\delta}.$$

and $\sigma^2_{\boldsymbol{\beta}}$ is the sum of the residual squared.

Next, I draw $\boldsymbol{\gamma}$. Let $\boldsymbol{\eta}$ be the vector whose j^{th} element is η_j. Then, $\boldsymbol{\gamma} \sim N\left(\widehat{\boldsymbol{\gamma}}, \sigma^2_{\boldsymbol{\gamma}} \left(\mathbf{X}'\mathbf{X}\right)^{-1}\right)$
, where

$$\widehat{\boldsymbol{\gamma}} = \left[\mathbf{X}'\mathbf{X}\right]^{-1} \mathbf{X}'\boldsymbol{\eta}.$$

Next, I draw $\boldsymbol{\mu_x}$. Let μ_{xk} be the kth element of $\boldsymbol{\mu_x}$. Draw μ_{xk} as

$$\mu_{xk} \sim N(\overline{log(x_k)}, \sigma^2_{xk}).$$

μ_ξ and μ_v are sampled as

$$\mu_\xi \sim N(\bar{\xi}, \sigma^2_\xi), \;\; \mu_v \sim N(\bar{v}, \sigma^2_v)$$

where ξ and v are retrieved from the demand and cost functions. Similarly, σ_{xk} that follows the log-normal distribution is drawn as follows:

$$\sigma^2_{xk} \sim IVG\left(m/2,\; \sum_k (log(x_{kj}) - \mu_{xk})^2/m\right)$$

σ_ξ, σ_v are drawn as:

$$\sigma^2_\xi \sim IVG\left(m/2,\; \xi'\xi/2m\right),\; \sigma^2_v \sim IVG\left(m/2,\; v'v/2m\right)$$

The rest of the parameters, $(EC, \sigma_c, \sigma_l, \rho)$ are estimated using the Metropolis-Hastings Algorithm.

Estimation Procedure The MCMC estimation procedure is summarized as follows.

1. Simulate $(\mathbf{x}, s, p)$ using the augmentation method.

2. For nonentrant nursing homes, derive the price and market share by solving their profit maximizing problem.

3. Draw the structural parameters. Use the Metropolos-Hastings algorithm to sample $(EC, \rho, \sigma_c, \sigma_l)$ and use the Gibbs sampling for the rest of the parameters.

4. Repeat the steps 1 to 3 10,000 times. Use the last 5,000 samples of the structural parameters to retrieve the posterior distribution of the parameters.

After all these parameters are estimated, I estimate κ using (82). For Medicaid patients' reimbursement rate, wage, and rental rate, I use the mean of the corresponding variables of entrants.

6.4 Results

Table 8 shows the estimation results. I report the mean and percentile statistics for the last 5,000 samples. My analysis has led to several noteworthy conclusions.

Firstly, the predicted mean own price elasticity, denoted by -4.097, closely mirrors the estimate of -4.56 as reported by Ching et al. (2015). This parallel between my findings and those of Ching et al. (2015) enhances the credibility of my estimation results, suggesting a robust validation of my methodological approach.

Secondly, I find that the aggregate of α_c and β_c amounts to 1.361, signaling increasing returns to scale. This finding is particularly relevant for policymaking, as it implies that government initiatives to invest more in nursing home facilities, concentrating investments in larger establishments proves more cost-efficient than constructing numerous smaller entities. This insight could significantly influence strategies aimed at optimizing healthcare infrastructure to accommodate the growing elderly population.

Lastly, my calculations indicate an estimated daily entry cost of \$10.9, or an annual figure of \$4,102, markedly lower in comparison to the total operational costs; an average nursing home incurs daily expenses of \$11,704, or about \$4.27 million annually. This discrepancy could be attributed to the heightened competition within the nursing home sector. The distribution of daily profits among nursing facilities, illustrated in Graph 13, provides further insight into this competitive landscape. The profit, calculated as the revenue from the PPP (price multiplied by the number of PPPs) less the total costs for the PPP (overall facility cost multiplied by the share of PPP within the facility), shows a predominance of low-profit scenarios across the industry. The median daily profit stands at merely \$13.61, with only 40.5% of nursing homes achieving daily profits that surpass the daily entry costs.

This situation suggests that the relatively low entry barrier encourages new firms to enter the market, thereby escalating competition. While a higher entry cost might deter potential entrants, allowing existing firms to enjoy higher profits, the current low barrier has led to an environment where the median profit is nearly zero. This dynamic highlights the critical role of entry costs in influencing market competition and the profitability landscape in the nursing home industry.

Selection Effect

Firms in this model are more likely to enter the market when their profits are higher. Consequently, nursing homes with higher ξ values and lower v values are expected to enter the market with a higher probability. To verify this, I estimated the kernel densities of ξ and v for entrants and compared them with the densities for all nursing homes. Figures 2 and 3 illustrate the distribution comparisons between entrants and all nursing homes.

As anticipated, entrant nursing homes exhibit higher average ξ values and lower average v values compared to the broader population of nursing homes. Additionally, the estimated densities for entrants demonstrate lower variance, a finding corroborated by the counterfactual simulations presented in Table 2. This evidence suggests that traditional instrumental variables may be invalid when dealing with the endogeneity.

Ching et al. (2015) utilizes the metric $\mathbf{x}_j\boldsymbol{\beta} + \xi_j$ to evaluate the quality of nursing homes. Figure 4 presents a comparison of the quality index between entrants and all nursing homes.[20] The average quality index of entrants is -4.82, compared to -5.34 for all potential entrants, indicating a 9.6% improvement. This comparison highlights the differences in quality measures and supports the findings regarding the market entry behavior of nursing homes.

6.5 Counterfactual Simulations

This subsection describes the results of counterfactual simulations. At the year of 2000, nursing homes could not increase their bed capacity due to the CON law. I simulated the market assuming there are no such restriction.[21] Under this counterfactual simulation, firm j's market share of MP is decided by the following equation:

$$s_j^m = \frac{\exp\left(\kappa(\mathbf{x}_j\boldsymbol{\beta} + \xi_j)\right)}{1 + \sum_{k \in J_e} \exp\left(\kappa(\mathbf{x}_k\boldsymbol{\beta} + \xi_k)\right)}. \tag{87}$$

Contrary to the market share with bed constraint, (72), the counterfactual market share is determined solely by the characteristics of the nursing homes. The following subsection describes the implications without the CON law.

[20]The distribution of the quality index for all nursing homes was derived by simulating $\mathbf{x}$ and ξ and subsequently estimating using the kernel density method.

[21]I did not simulate firms' entry decisions because bed constraints do not affect firms' entry decisions; firms in this model decide based on the profit level of PPP. Therefore, all I did in this counterfactual was to recalculate the entrant's profit with a modified market share and calculate the social welfare.

6.5.1 Effect of the CON law on Social Welfare

First I calculate the consumer surplus following Ching et al. (2015). The consumer i's surplus under the bed restriction is given as

$$E[CS_{i,r}^m] = \frac{1}{\alpha} \log \left(\sum_{j \in J_r} \exp(u_j^m) \right). \quad (88)$$

where the total consumer surplus is given as

$$TCS^m = \sum_r M_r E[CS_{i,r}^m]. \quad (89)$$

Under the counterfactual, the consumer i's surplus is given by

$$E[CS_i^m] = \frac{1}{\alpha} E\left[\max(u_{ij}^m)\right] = \frac{1}{\alpha} \log \left(\sum_{j \in J} \exp(\bar{u}_j^m) \right), \quad (90)$$

where $\bar{u}_j^m$ is the utility perceived from nursing home j's characteristics[22]. The total consumer surplus can be expressed as

$$TCS^m = M^m E[CS_i^m] \quad (91)$$

Next, the producer surplus is calculated by summing the difference between the reimbursement rate and the marginal cost for each nursing home.

$$PS = \sum_j^{286} \sum_i^{q_{mj}} \left(p_j^m - MC_j(i)\right) - EC \quad (92)$$

The daily and annual consumer surplus increased by \$2.46 million and \$901 million, respectively. The producer surplus increased by \$444,821 daily and \$162 million annually. Government expenditure rose by \$1.88 million daily and \$689 million annually. Overall, the counterfactual social welfare increased by \$1.02 million daily or \$374.3 million annually.

7 Conclusion

This paper introduces a novel instrument-free approach for estimating structural parameters in differentiated products markets, specifically in cases where traditional instruments fail due

[22] I assume that MP and PPP have the same sensitivity towards price, i.e. price coefficient.

to endogenous entry decisions. The theoretical framework developed in this paper extends the logit and random coefficient logit models of demand and cost estimation by incorporating firms' entry-exit decisions into the market. While existing studies like Berry et al. (1995) provide tools for estimating demand and supply parameters under static market conditions, they often ignore the complexities introduced by firms' strategic decisions to enter or exit a market. This oversight leads to biased estimates, especially when firm-level data for non-entrants is unavailable, as is often the case in industries with significant regulatory barriers like nursing homes.

The theoretical contribution of this paper lies in addressing these limitations by developing an instrument-free estimation strategy, building on the recent methodologies proposed by Byrne et al. (2022) and Imai et al. (2024). By utilizing available cost data and entrant firms' profit functions, this approach effectively circumvents the endogeneity issues without relying on traditional instruments such as cost shifters or rival firms' characteristics. Through Monte Carlo simulations, I demonstrate that this approach yields consistent parameter estimates in oligopolistic markets, even in the absence of data for non-entrant firms. This theoretical innovation significantly enhances the robustness of structural estimations in markets where firm entry decisions are crucial but difficult to observe comprehensively.

Empirically, the paper applies this methodology to the nursing home industry in Wisconsin, a market heavily regulated by the CON law. The CON law, designed to control costs by limiting the expansion of healthcare facilities, has been widely criticized for creating excess demand and limiting access to care, particularly for Medicaid patients. By applying the structural estimation techniques developed in this paper, I quantify the impact of the CON law on market outcomes, including social welfare, government spending, and the quality of care in SNFs.

The empirical results show that the repeal of the CON law would generate substantial welfare gains. Specifically, the removal of entry restrictions would motivate 17,287 additional patients to access nursing home care annually, leading to an increase in consumer surplus of $868 million and producer surplus of $165 million. On the other hand, government expenditure is projected to increase by $700 million due to higher Medicaid reimbursement rates and increased utilization of services. However, despite this rise in government costs, the overall increase in social welfare amounts to $333 million per year, highlighting the net positive impact of removing the CON law on society.

Another key finding from the empirical analysis is the effect of competition on nursing home quality. The quality measure used in the study, based on a structural model of private-pay patients' choices, indicates that the quality of care provided by current market entrants is approximately 9.6% higher than that of potential entrants. This suggests that market competition driven by entry-exit dynamics plays a crucial role in improving service quality. The results align with previous literature, such as Hackmann (2019), which emphasizes the importance of competition in raising the standard of care in healthcare markets.

Moreover, the estimation results reveal that the aggregate of the cost parameters α_c and β_c amounts to 1.361, signaling increasing returns to scale in the nursing home industry. Notably, this study is the first to estimate these cost parameters while taking into account output endogeneity, addressing potential biases that arise when output levels are correlated with unobserved cost factors. By accounting for output endogeneity, the analysis provides more accurate and reliable estimates of the cost structure. This finding is particularly relevant for policymaking, as it implies that government initiatives to invest more in nursing home facilities should concentrate investments in larger establishments, which proves more cost-efficient than constructing numerous smaller entities. This insight could significantly influence strategies aimed at optimizing healthcare infrastructure to accommodate the growing elderly population.

Despite these contributions, the study acknowledges certain limitations that present avenues for future research. First, while econometricians do not need to observe non-entrants' product characteristics, they still need to observe the number of non-entrants to correctly identify the parameters, which is a strong assumption. Therefore, the developed model is primarily applicable in markets where researchers know which companies are considering entering the market but do not have information about their products. Second, the supply side of the nursing home industry is oversimplified due to the assumption of perfect competition in the labor market. However, as reported in Nevidjon and Erickson (2001), the industry experiences a shortage of nurses, invalidating the perfect competition assumption. Indeed, as shown in Table 11, nurse wages vary across nursing homes, suggesting imperfect competition.

In conclusion, this paper makes significant contributions to both the theoretical and empirical literature on market entry and competition by developing and applying a new estimation strategy that effectively addresses the complexities of entry-exit decisions. The findings provide strong evidence that easing regulatory barriers, such as the CON law, can substantially improve market efficiency, increase social welfare, and enhance the quality of essential services. Future research can build on this work by applying the instrument-free estimation methodology to other regulated industries, thereby offering further insights into the broader effects of entry regulation and competition on market outcomes.

References

Aguirregabiria, V., A. Iaria, and S. Sokullu (2024): "Identification and Estimation of Demand Models with Endogenous Product Entry and Exit," Tech. rep., Working Paper.

Amemiya, T. (1973): "Regression analysis when the dependent variable is truncated normal," *Econometrica: Journal of the Econometric Society*, 997–1016.

Bardey, D. and L. Siciliani (2021): "Nursing-homes' competition and distributional implications when the market is two-sided," *Journal of economics & management strategy*, 30, 472–500.

Bekele, G. and A. Holtmann (1987): "A cost function for nursing homes: Toward a system of diagnostic reimbursement groupings," *Eastern Economic Journal*, 13, 115–122.

Berry, S. T. (1992): "Estimation of a Model of Entry in the Airline Industry," *Econometrica: Journal of the Econometric Society*, 889–917.

——— (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242–262.

Berry, S. T. and P. Haile (2014): "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, forthcoming.

Berry, S. T., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841–890.

Bontemps, C., C. Hualdani, and K. Remmy (2023): "Price Competition and Endogenous Product Choice in Networks: Evidence from the US Airline Industry," *Working Paper.*

Bresnahan, T. F. and P. C. Reiss (1990): "Entry in Monopoly Markets," *Review of Economic Studies*, 57, 531–553.

——— (1991): "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99, 977–1009.

Bresnahan, T. F., P. C. Reiss, R. Willig, and G. J. Stigler (1987): "Do entry conditions vary across markets?" *Brookings Papers on Economic Activity*, 1987, 833–881.

Byrne, D. P., S. Imai, N. Jain, and V. Sarafides (2022): "Identification and Estimation of Differentiated Products Models without Instruments using Cost data," *Journal of Econometrics*, 228(2), 278–301.

Ching, A. T., F. Hayashi, and H. Wang (2015): "Quantifying the impacts of limited supply: The case of nursing homes," *International Economic Review*, 56, 1291–1322.

Christensen, E. W. (2004): "Scale and scope economies in nursing homes: a quantile regression approach," *Health economics*, 13, 363–377.

Ciliberto, F., C. Murry, and E. Tamer (2021): "Market structure and competition in airline markets," *Journal of Political Economy*, 129, 2995–3038.

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

Cohen, J. W. and W. D. Spector (1996): "The effect of Medicaid reimbursement on quality of care in nursing homes," *Journal of health economics*, 15, 23–48.

Draganska, M., M. Mazzeo, and K. Seim (2009): "Beyond plain vanilla: Modeling joint product assortment and pricing decisions," *QME*, 7, 105–146.

Dudzinski, C. S., O. Homer Erekson, and A. L. Ziegert (1998): "Estimating an hedonic translog cost function for the home health care industry," *Applied Economics*, 30, 1259–1267.

Duncan, G. M. (1987): "A simplified approach to M-estimation with application to two-stage estimators," *Journal of Econometrics*, 34, 373–389.

Eizenberg, A. (2014): "Upstream innovation and product variety in the us home pc market," *Review of Economic Studies*, 81, 1003–1045.

Ericson, R. and A. Pakes (1995): "Markov-perfect industry dynamics: A framework for empirical work," *The Review of economic studies*, 62, 53–82.

Ettner, S. L. (1993): "Do elderly Medicaid patients experience reduced access to nursing home care?" *Journal of Health Economics*, 12, 259–280.

Farronato, C. and A. Fradkin (2022): "The welfare effects of peer entry: the case of Airbnb and the accommodation industry," *American Economic Review*, 112, 1782–1817.

Filippini, M. (2001): "Economies of scale in the Swiss nursing home industry," *Applied Economics Letters*, 8, 43–46.

Gertler, P. J. (1989): "Subsidies, quality, and the regulation of nursing homes," *Journal of Public Economics*, 38, 33–52.

——— (1992): "Medicaid and the Cost of Improving Access to Nursing Home Care," *The Review of Economics and Statistics*, 74, 338–45.

Gertler, P. J. and D. M. Waldman (1992): "Quality-adjusted cost functions and policy evaluation in the nursing home industry," *Journal of Political Economy*, 100, 1232–1256.

Giorgio, L. D., M. Filippini, and G. Masiero (2016): "Is higher nursing home quality more costly?" *The European Journal of Health Economics*, 17, 1011–1026.

Gowrisankaran, G. and R. J. Town (1997): "Dynamic equilibrium in the hospital industry," *Journal of Economics & Management Strategy*, 6, 45–74.

Grabowski, D. C. (2001): "Medicaid reimbursement and the quality of nursing home care," *Journal of health economics*, 20, 549–569.

Grabowski, D. C. and J. J. Angelelli (2004): "The relationship of medicaid payment rates, bed constraint policies, and risk-adjusted pressure ulcers," *Health Services Research*, 39, 793–812.

Grabowski, D. C., J. Gruber, and J. J. Angelelli (2008): "Nursing home quality as a common good," *The review of economics and statistics*, 90, 754–764.

Grant, I., I. Kesternich, and J. Van Biesebroeck (2022): "Entry decisions and asymmetric competition between non-profit and for-profit homes in the long-term care market," *International Economic Review*, 63, 631–670.

Gruenberg, L. W. and T. R. Willemain (1982): "Hospital discharge queues in Massachusetts," *Medical Care*, 20, 188–201.

Gupta, A., S. T. Howell, C. Yannelis, and A. Gupta (2024): "Owner incentives and performance in healthcare: Private equity investment in nursing homes," *The Review of Financial Studies*, 37, 1029–1077.

Hackmann, M. B. (2019): "Incentivizing better quality of care: The role of Medicaid and competition in the nursing home industry," *American Economic Review*, 109, 1684–1716.

Harrington, C., D. Zimmerman, S. L. Karon, J. Robinson, and P. Beutel (2000): "Nursing home staffing and its relationship to deficiencies," *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 55, S278–S287.

He, D., P. McHenry, and J. M. Mellor (2020): "The effects of Medicare payment changes on nursing home staffing," *American Journal of Health Economics*, 6, 411–443.

Heckman, J. J. (1976): "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," in *Annals of economic and social measurement, volume 5, number 4*, NBER, 475–492.

Heger, D., A. Herr, and A. Mensen (2022): "Paying for the view? How nursing home prices affect certified staffing ratios," *Health Economics*, 31, 1618–1632.

Ho, K. (2009): "Insurer-provider networks in the medical care market," *American Economic Review*, 99, 393–430.

Hofler, R. A. and B. Rungeling (1994): "US nursing homes: are they cost efficient?" *Economics Letters*, 44, 301–305.

Horn, S. D., P. Buerhaus, N. Bergstrom, and R. J. Smout (2005): "RN staffing time and outcomes of long-stay nursing home residents: pressure ulcers and other adverse outcomes are less likely as RNs spend more time on direct patient care." *AJN The American Journal of Nursing*, 105, 58–70.

Imai, S., J. Neelam, H. Suzuki, and M. Taniguchi (2024): "Estimating Cost Functions in Differentiated Product Oligopoly Models without Instruments," *Working Paper*.

Jia, P. (2008): "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry," *Econometrica*, 76, 1263–1316.

Jiang, R., P. Manchanda, and P. E. Rossi (2009): "Bayesian analysis of random coefficient logit models using aggregate data," *Journal of Econometrics*, 149, 136–148.

Knox, K. J., E. C. Blankmeyer, and J. Stutzman (2004): "Administrative compensation and organizational performance in Texas nursing facilities," *Small Business Economics*, 22, 33–49.

Kotschy, R. and D. E. Bloom (2022): "A comparative perspective on long-term care systems," *International Social Security Review*, 75, 47–69.

Kuehn, J. (2018): "Spillovers from entry: the impact of bank branch network expansion," *The RAND Journal of Economics*, 49, 964–994.

Kunz, J. S., C. Propper, K. E. Staub, and R. Winkelmann (2024): "Assessing the quality of public services: For-profits, chains, and concentration in the hospital market," *Health Economics*.

Li, S., J. Mazur, Y. Park, J. Roberts, A. Sweeting, and J. Zhang (2022): "Repositioning and market power after airline mergers," *The RAND Journal of Economics*, 53, 166–199.

Lin, H. (2015): "Quality choice and market structure: A dynamic analysis of nursing home oligopolies," *International Economic Review*, 56, 1261–1290.

Lu, S. F., K. Serfes, G. Wedig, and B. Wu (2021): "Does competition improve service quality? The case of nursing homes where public and private payers coexist," *Management Science*, 67, 6493–6512.

Mazzeo, M. J. (2002): "Product choice and oligopoly market structure," *RAND Journal of Economics*, 33, 221–242.

Melitz, M. J. (2003): "The impact of trade on intra-industry reallocations and aggregate industry productivity," *econometrica*, 71, 1695–1725.

Miller, N., S. Berry, F. Scott Morton, J. Baker, T. Bresnahan, M. Gaynor, R. Gilbert, G. Hay, G. Jin, B. Kobayashi, et al. (2022): "On the misuse of regressions of price on the HHI in merger review," *Journal of Antitrust Enforcement*, 10, 248–259.

Miller, N. H. and M. C. Weinberg (2017): "Understanding the price effects of the MillerCoors joint venture," *Econometrica*, 85, 1763–1791.

Mizuno, T. (2003): "On the existence of a unique price equilibrium for models of product differentiation," *International Journal of Industrial Organization*, 21, 761–793.

Murmann, M., D. Sinden, A. T. Hsu, K. Thavorn, A. B. Eddeen, A. H. Sun, and B. Robert (2023): "The cost-effectiveness of a nursing home-based transitional care unit for increasing the potential for independent living in the community among hospitalized older adults," *Journal of Medical Economics*, 26, 61–69.

Mutter, R. L., W. H. Greene, W. Spector, M. D. Rosko, and D. B. Mukamel (2013): "Investigating the impact of endogeneity on inefficiency estimates in the application of stochastic frontier analysis to nursing homes," *Journal of Productivity Analysis*, 39, 101–110.

Nevidjon, B. and J. I. Erickson (2001): "The nursing shortage: Solutions for the short and long term," *Online Journal of Issues in Nursing*, 6.

Nevo, A. (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307–342.

Norton, E. C. (1992): "Incentive regulation of nursing homes," *Journal of Health Economics*, 11, 105–128.

Nyman, J. A. (1985): "Prospective and 'cost-plus' Medicaid reimbursement, excess Medicaid demand, and the quality of nursing home care," *Journal of Health Economics*, 4, 237–259.

——— (1988a): "The effect of competition on nursing home expenditures under prospective reimbursement." *Health Services Research*, 23, 555.

——— (1988b): "Excess demand, the percentage of Medicaid patients, and the quality of nursing home care," *Journal of Human Resources*, 76–92.

——— (1994): "The effects of market concentration and excess demand on the price of nursing home care," *The Journal of Industrial Economics*, 193–204.

Park, Y. (2020): "Structural remedies in network industries: An assessment of slot divestitures in the American Airlines/US Airways merger," *Working Paper*.

Petrin, A. (2002): "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, 110, 705–729.

Reiss, P. C. and P. T. Spiller (1989): "Competition and entry in small airline markets," *The Journal of Law and Economics*, 32, S179–S202.

Seim, K. (2006): "An empirical model of firm entry with endogenous product-type choices," *The RAND Journal of Economics*, 37, 619–640.

Sweeting, A. (2013): "Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81, 1763–1803.

Tanaka, M. (2020): "Adaptive MCMC for generalized method of moments with many moment conditions," *ISC2019*.

Tobin, J. (1958): "Estimation of relationships for limited dependent variables," *Econometrica: journal of the Econometric Society*, 24–36.

Vitaliano, D. F. and M. Toren (1994): "Cost and efficiency in nursing homes: a stochastic frontier approach," *Journal of Health economics*, 13, 281–300.

Wollmann, T. G. (2018): "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles," *American Economic Review*, 108, 1364–1406.

Yang, O., J. Yong, and A. Scott (2022): "Nursing home competition, prices, and quality: a scoping review and policy lessons," *The Gerontologist*, 62, e384–e401.

Yuan, Z. and P. Jia (2024): "Network Competition in the Airline Industry: An Empirical Framework," *NBER Working Paper*.

8 Tables and Figures

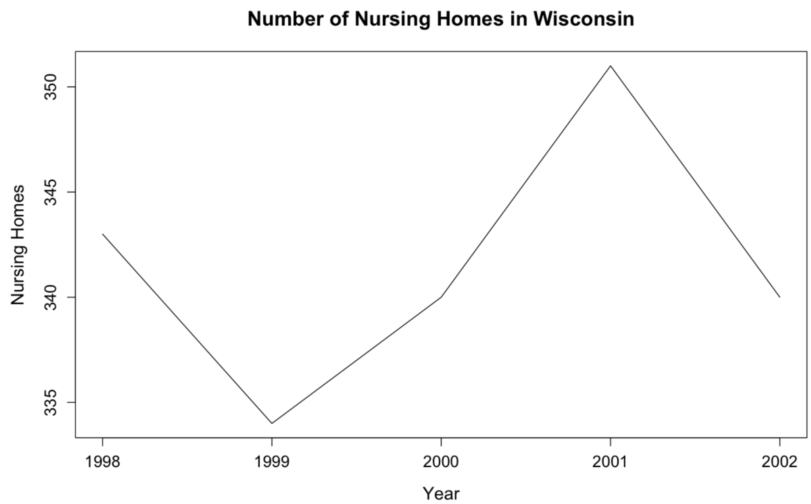

Figure 1: Trends in the Number of Nursing Homes

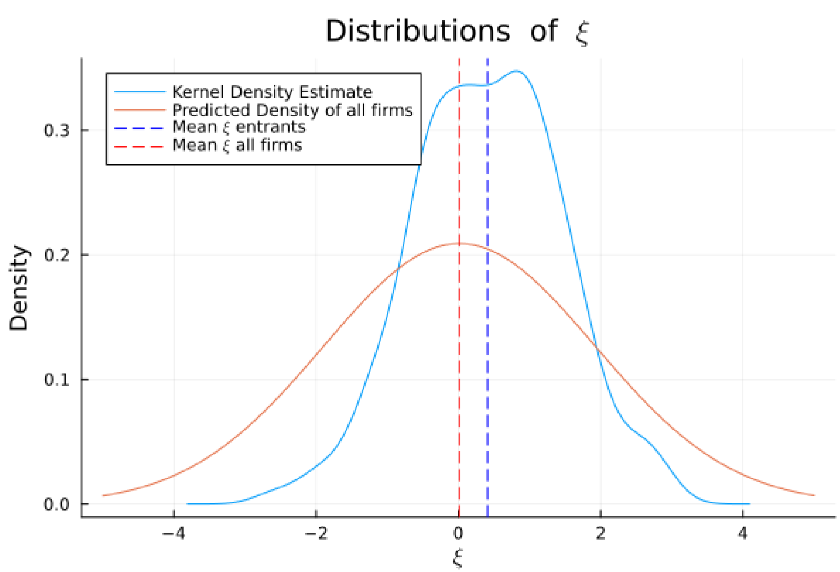

Figure 2: Entrants' and all nursing homes' ξ

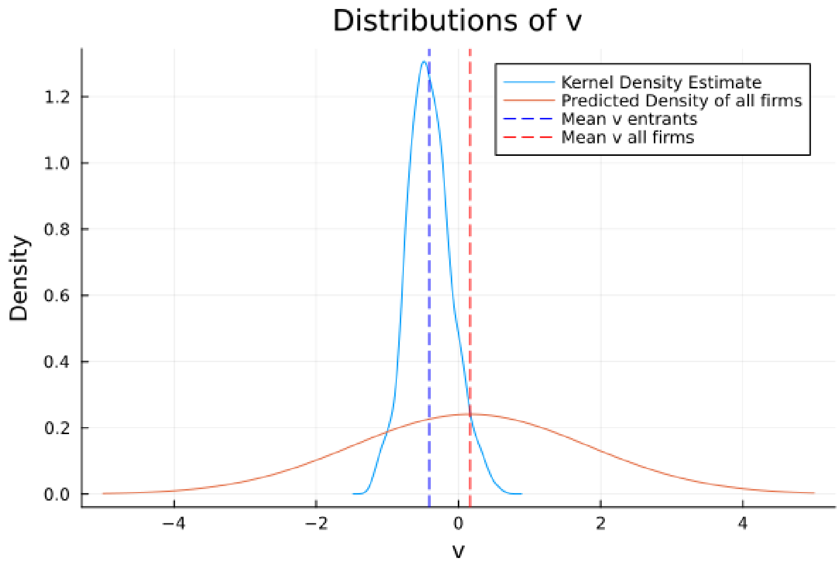

Figure 3: Entrants' and all nursing homes' v

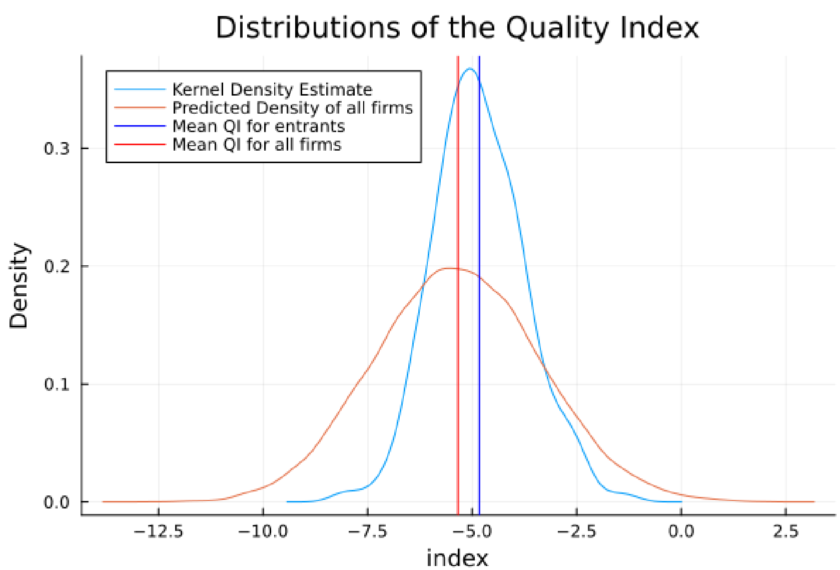

Figure 4: Distributions of the quality index

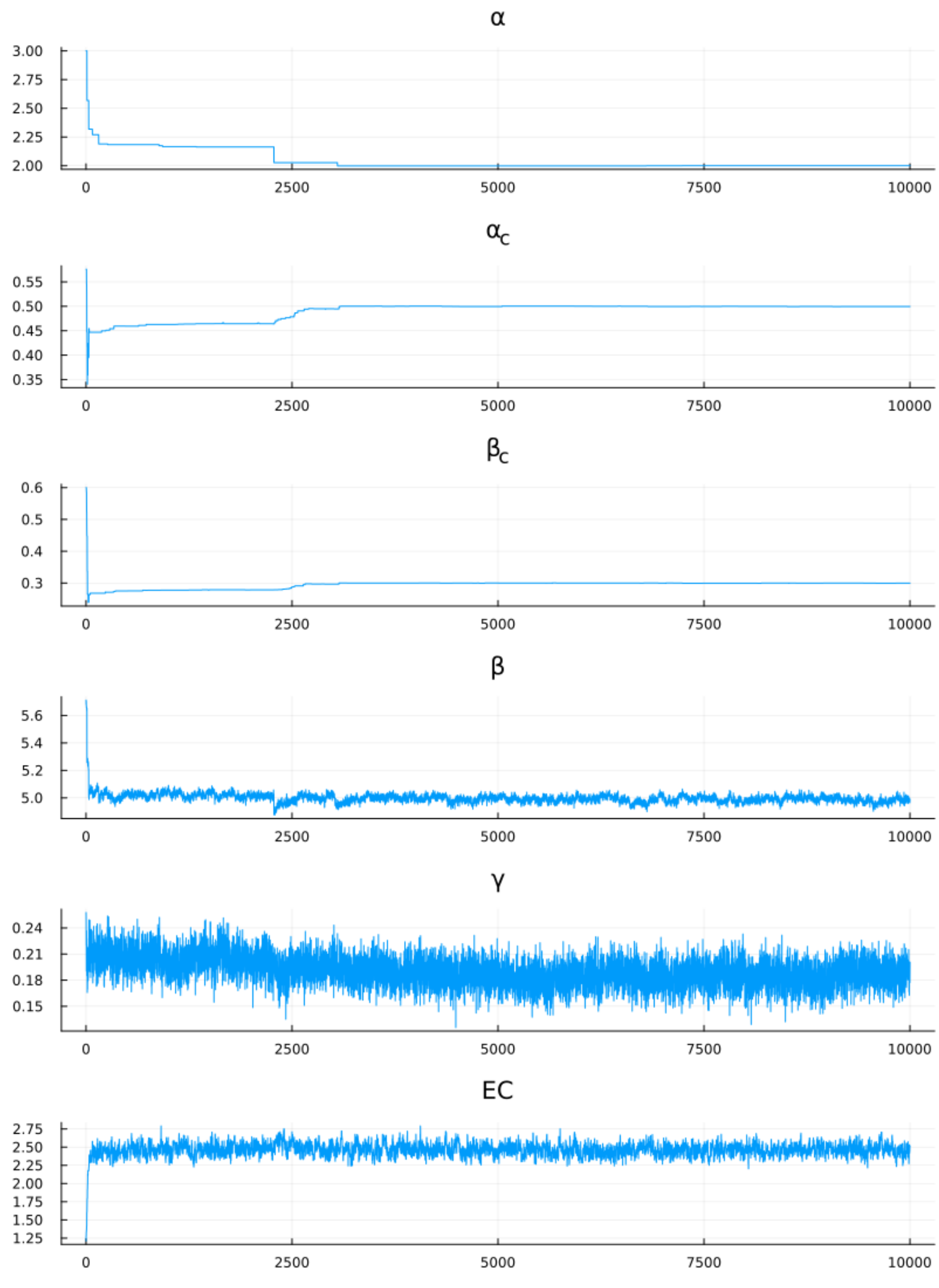

Figure 5: Monte-Carlo: MCMC Samples from the Monopoly Model

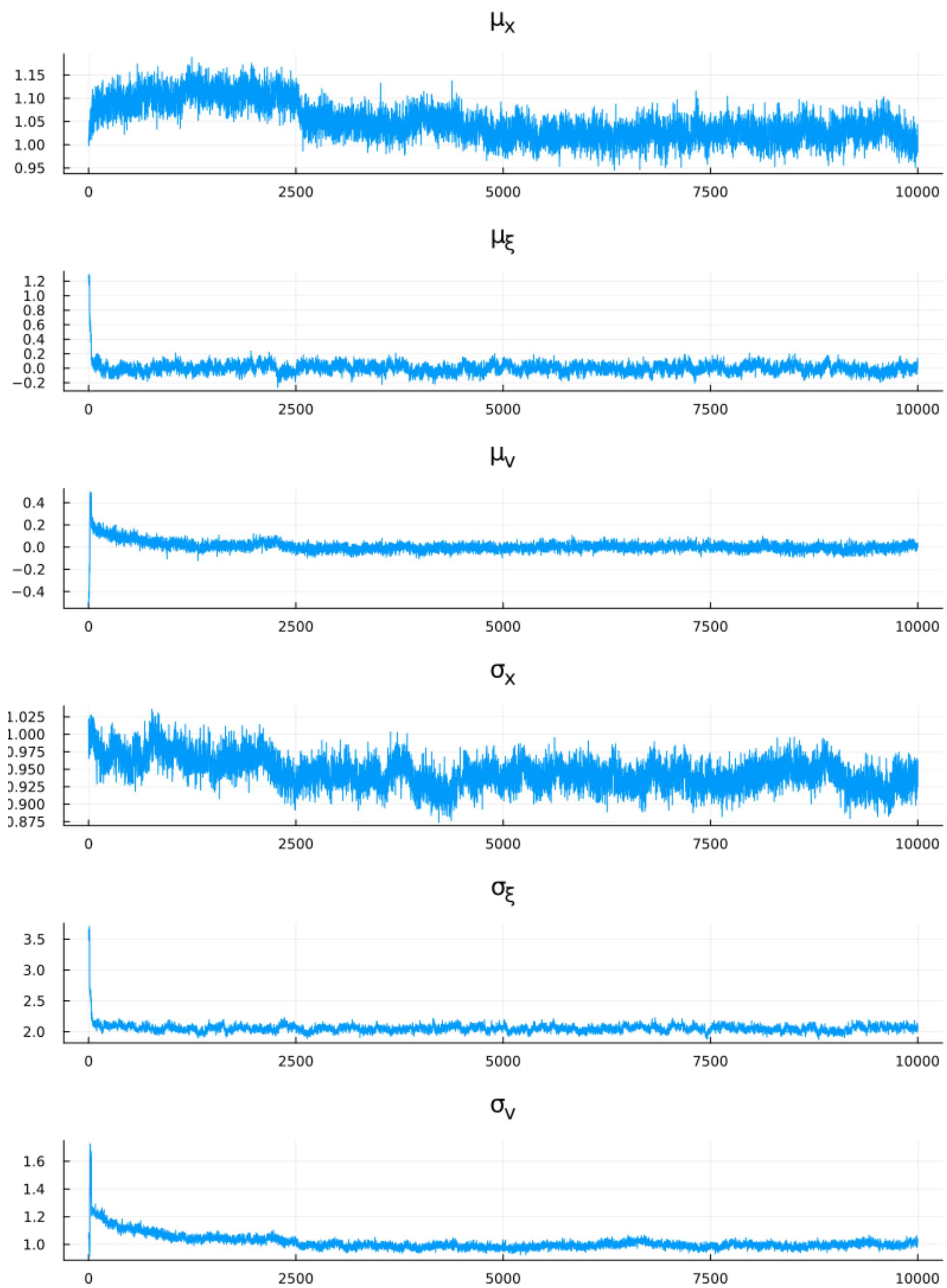

Figure 6: Monte-Carlo: MCMC Samples from the Monopoly Model

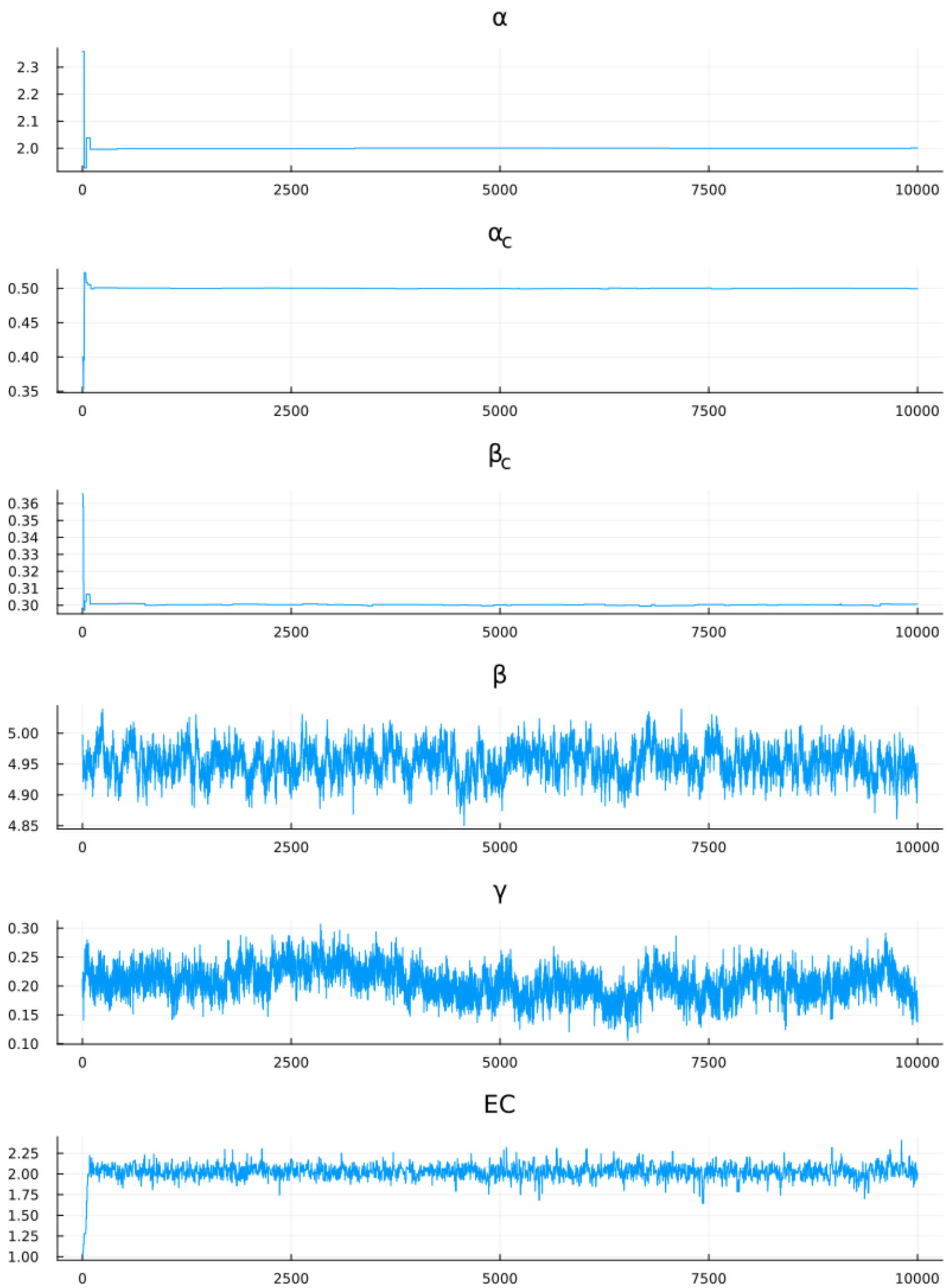

Figure 7: Monte-Carlo: MCMC Samples from Duopoly Model

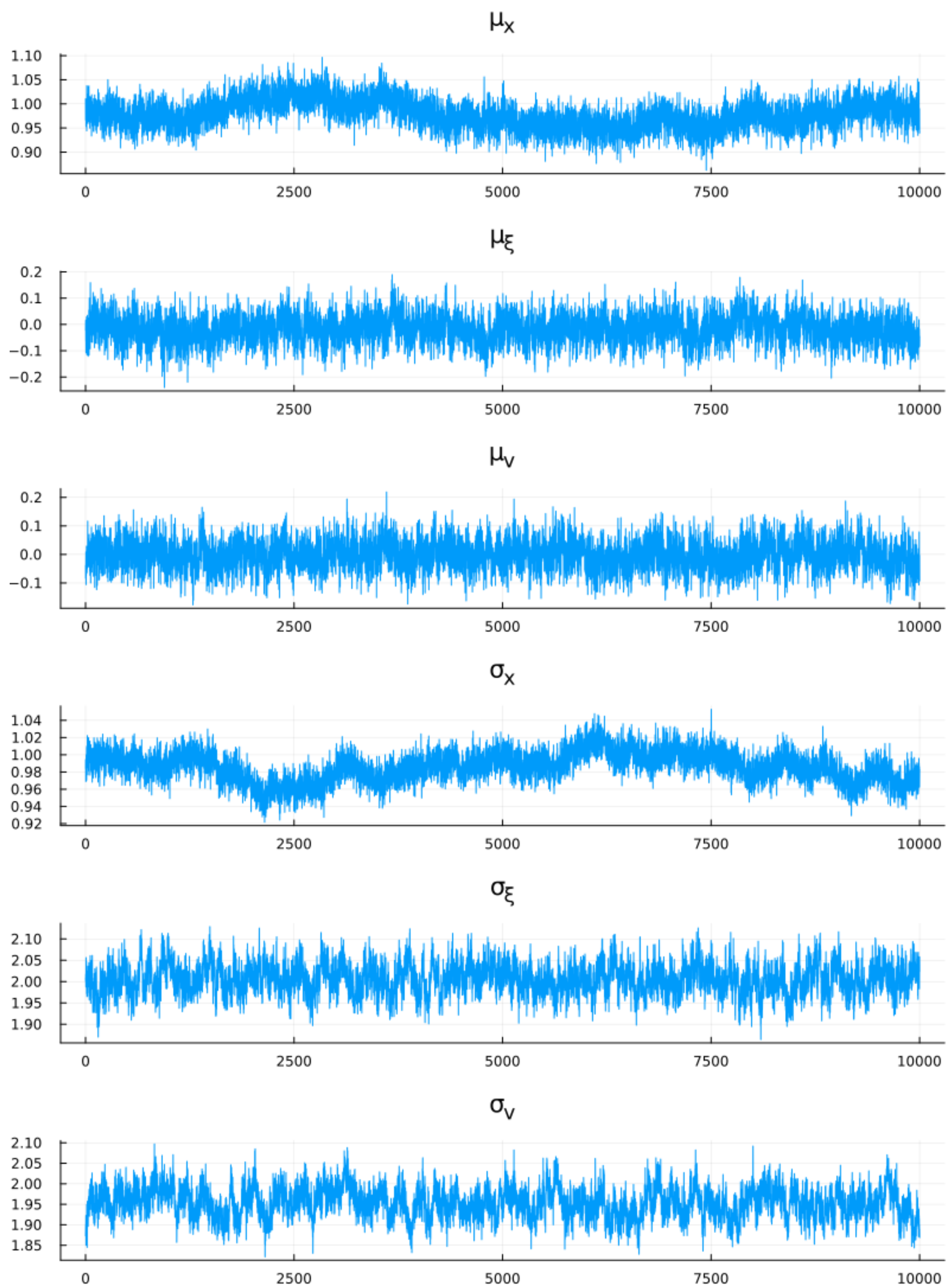

Figure 8: Monte-Carlo: MCMC Samples from Duopoly Model

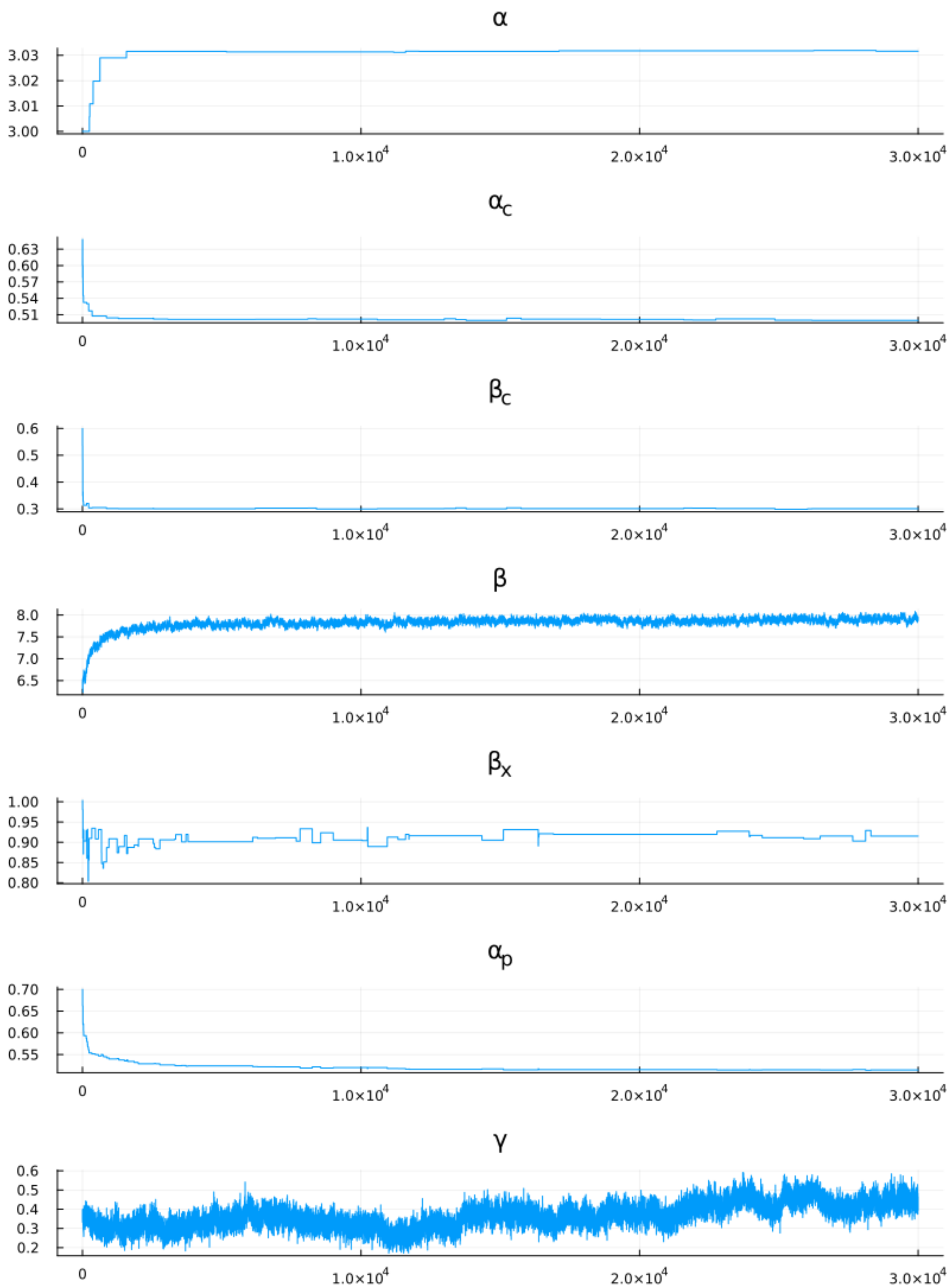

Figure 9: Monte-Carlo: MCMC Samples from the BLP Model

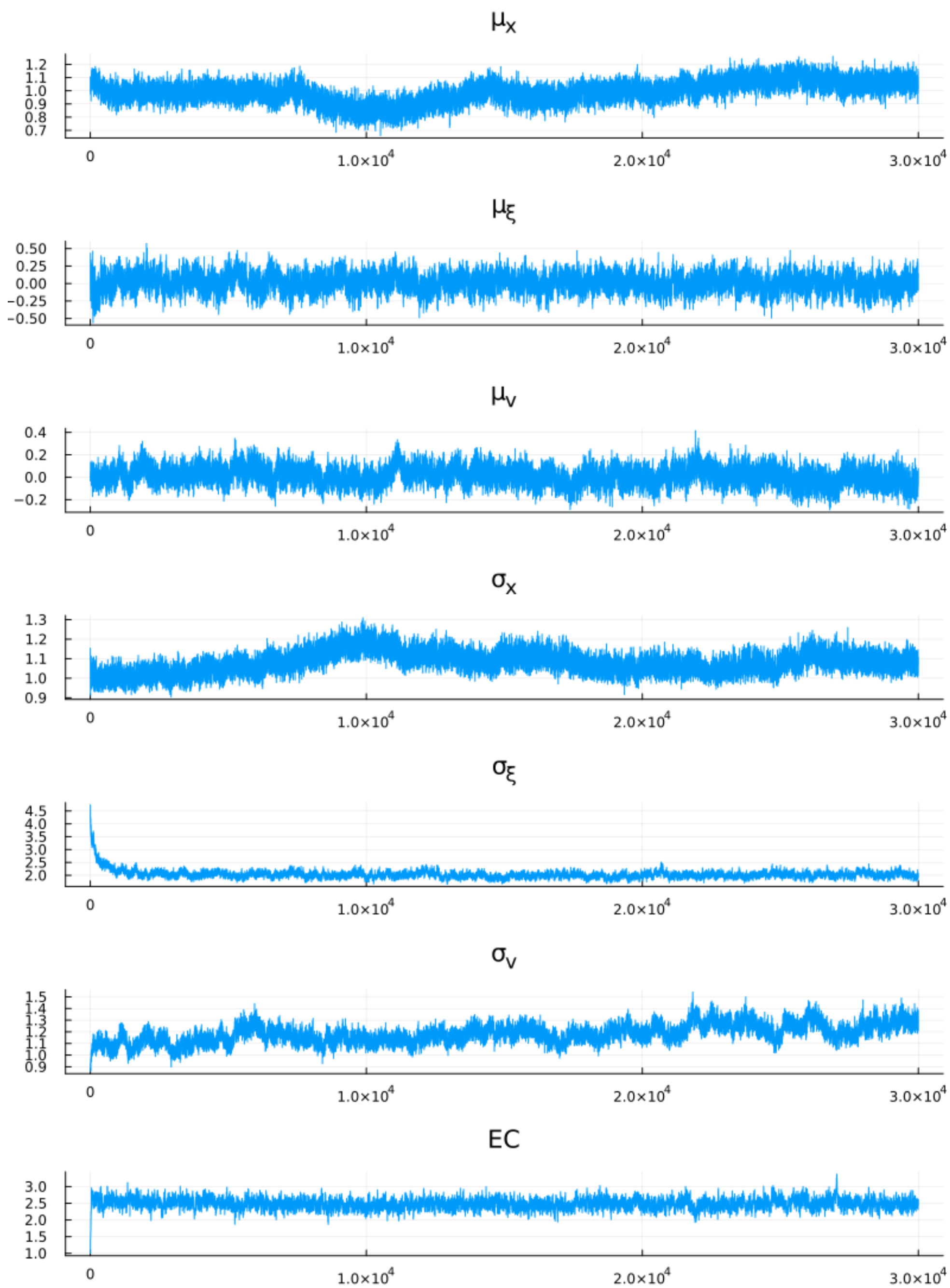

Figure 10: Monte-Carlo: MCMC Samples from the BLP Model

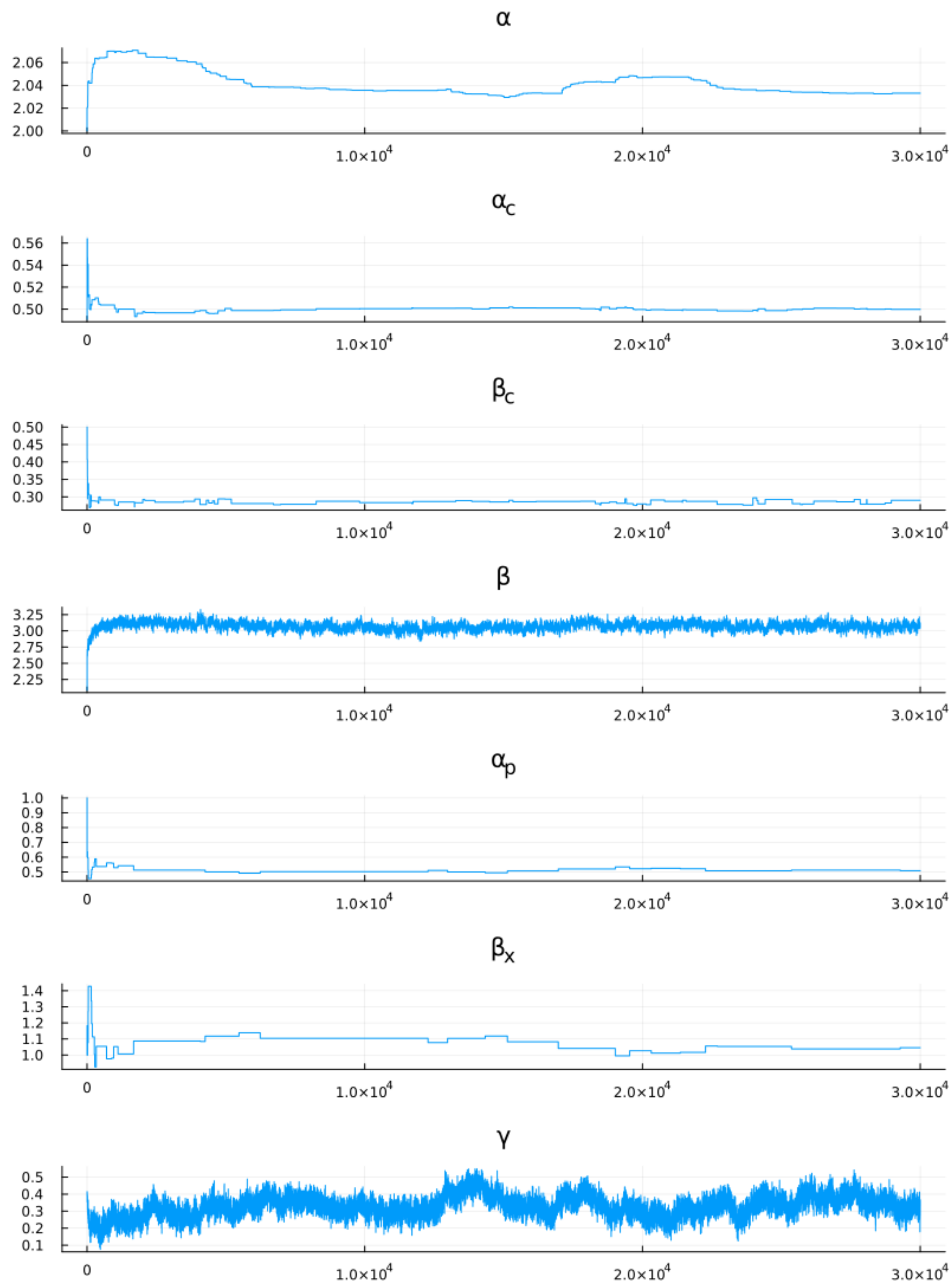

Figure 11: Monte-Carlo: MCMC Samples from the BLP Model with Moment Conditions

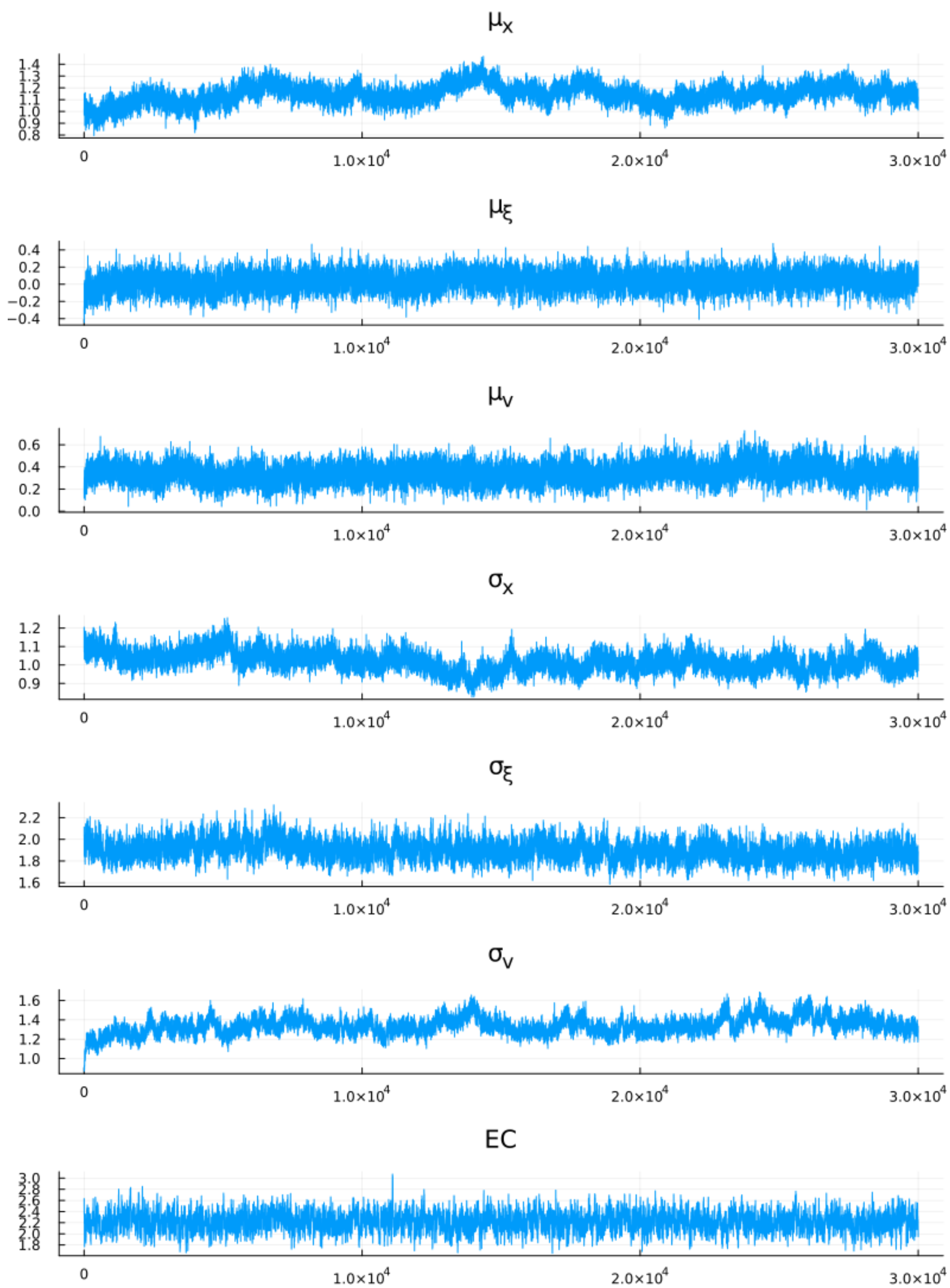

Figure 12: Monte-Carlo: MCMC Samples from the BLP Model with Moment Conditions

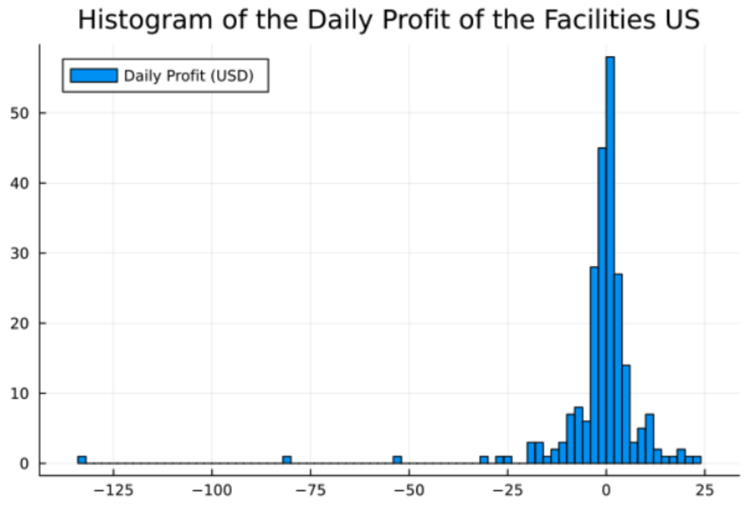

Figure 13: Histogram of Daily Profit of Nursing Homes

Table 1: Monte-Carlo Parameter Values

Parameter	Description	Value
(a) Demand-side parameters		
α	Price coef	2.0
β	Product characteristic coef. mean	5.0
γ	Product characteristic coef. mean	0.2
μ_x	Product characteristic mean	1.0
σ_x	Product characteristic std. dev.	1.0
μ_ξ	Unobserved product quality mean	0.0
σ_ξ	Unobserved product quality std. dev.	2
Q_L	Lower bound on market size	10
Q_H	Upper bound on market size	15
(b) Supply-side parameters		
α_c	Labor coef. in Cobb-Douglas prod. fun.	0.5
β_c	Capital coef. in Cobb-Douglas prod. fun.	0.3
μ_w	log wage mean	0.0
σ_w	log wage std. dev.	1.0
μ_r	log rental rate mean	0.0
σ_r	Rental rate std. dev.	1.0
μ_v	log cost shock mean	0.0
σ_v	Cost shock std. dev.	1.0
E	Entry cost	2.5
J	Number of firms in each market	1
(c) Cost measurement error		
σ_c	Measurement std. dev., cost	0.15
σ_l	Measurement std. dev., labor cost	0.1
ρ	Correlation between the measurement errors	0.5

Table 2: Sample Statistics

A. Entrants and Non-entrants				
Variable	True Mean	Sample Mean	True Std. Dev	Sample Std. Dev.
x^*	1.0	0.997	1.225	1.225
lnr^*	0	0.002	1.0	1.001
lnw^*	0	0.001	1.0	0.999
ξ^*	0	0.004	1.5	1.502
υ^*	0	0.002	1.0	1.000
Price		4.109		2.196
Quantity		3.067		3.985

B. Entrants				
Variable	True Mean	Sample Mean	True Std. Dev	Sample Std. Dev.
x	1.0	1.297	1.225	1.192
lnr	0	-0.174	1.0	0.987
lnw	0	-0.265	1.0	1.451
ξ	0	0.679	1.5	1.289
υ	0	-0.369	1.0	0.881
Price		4.624		2.057
Quantity		6.462		3.844

C. Correlation				
	Entrants and Non-Entrants		Entrants	
	ξ	υ	ξ	υ
x	0.0001	-0.002	-0.143	0.073
r	0.002	0.002	0.058	-0.06
w	0.001	-0.003	0.078	-0.131
ξ	1.0	-0.002	1.0	0.073
Price	0.2764	0.4368	0.2467	0.4145
Quantity	0.1795	-0.4660	0.13268	-0.4337

Table 3: Demand and cost IV estimates (sd in parenthesis)

		(1)	(2)	(3)
Parameters	True	OLS Estimate (all)	IV (all)	IV (Entrants)
α	2	1.368 (0.028)	1.999 (0.080)	1.897 (0.192)
β	5	4.115 (0.049)	4.999 (0.118)	4.653 (0.333)
α_c	0.5	0.702 (0.021)	0.497 (0.019)	0.883 (0.107)
β_c	0.3	0.438 (0.019)	0.303 (0.015)	0.554 (0.160)
γ	0.2	0.243 (0.011)	0.199 (0.013)	0.292 (0.044)

# of markets: 2000.

The figures shown in table 3 and 4 are the mean of the 100 simulated results.

Variables	True	MCMC (mean)	sd
α	2	2.000	0.0005
α_c	0.5	0.499	0.0002
β_c	0.3	0.299	0.0002

Table 4: IV-Free OLS estimation

Variables	True	MCMC (mean)	sd
α	2	1.999	0.0003
α_c	0.5	0.5002	0.0001
β_c	0.3	0.3000	0.0002
β	5	4.986	0.023
γ	0.2	0.194	0.014
EC	2.5	2.517	0.085
x_{mean}	1	1.018	0.024
ξ_{mean}	0	-0.035	0.060
v_{mean}	0	0.011	0.029
σ_x	1	0.969	0.021
σ_ξ	2	2.01	0.053
σ_v	1	1.015	0.019
ρ	0.5	0.491	0.025
σ_l	0.1	0.097	0.002
σ_c	0.15	0.147	0.003

Table 5: Monte-Carlo result: MCMC Posterior Distributions, Monopoly

Variables	True	MCMC (mean)	sd
α	2.0	1.99	0.009
α_c	0.5	0.504	0.015
β_c	0.3	0.298	0.015
β	5.0	4.984	0.023
γ	0.2	0.195	0.015
EC	2.5	2.515	0.087
x_{mean}	1	1.023	0.023
ξ_{mean}	0	-0.026	0.062
v_{mean}	0	0.021	0.031
σ_x	1	0.947	0.018
σ_ξ	2	2.043	0.048
σ_v	1	1.010	0.022

Table 6: Monte-Carlo result: MCMC Posterior Distributions Without Cost Data

Variables	True	MCMC (mean)	sd
α	2.0	2.105	0.0002
α_c	0.5	0.436	0.030
β_c	0.3	0.315	0.031
β	5.0	4.964	0.017
γ	0.2	-0.005	0.014
EC	2.5	2.898	0.106
x_{mean}	1	1.195	0.055
ξ_{mean}	0	0.012	0.078
v_{mean}	0	-0.099	0.069
σ_x	3	2.239	0.038
σ_ξ	2	2.054	0.051
σ_v	3	2.347	0.045

Table 7: Monte-Carlo result: MCMC Posterior Distributions Without Cost Data, high variance

Variables	True	MCMC (mean)	sd
α	2.0	1.999	1.1e-15
α_c	0.5	0.500	0.0002
β_c	0.3	0.300	0.0002
β	5.0	4.952	0.0262
γ	0.2	0.218	0.024
EC	2.0	2.027	0.073
x_{mean}	1	0.996	0.0255
ξ_{mean}	0	-0.018	0.0515
v_{mean}	0	0.006	0.0513
σ_x	1	0.968	0.0139
σ_ξ	2	2.008	0.0367
σ_v	2	1.958	0.0350

Table 8: Monte-Carlo Simulation, Duopoly

Variables	True	MCMC (mean)	sd
α	3.0	3.032	0.000
α_c	0.5	0.501	0.001
β_c	0.3	0.301	0.001
β	8.0	7.881	0.056
β_x	0.9	0.919	0.006
α_p	0.5	0.515	0.000
γ	0.2	0.405	0.056
EC	2.5	2.484	0.165
x_{mean}	1	1.028	0.071
ξ_{mean}	0	0.006	0.120
v_{mean}	0	-0.003	0.000
σ_x	1	1.077	0.044
σ_ξ	2	2.003	0.101
σ_v	1	1.221	0.074

Table 9: Monte-Carlo Simulation: MCMC Posterior Distributions, BLP model

Variables	True	MCMC (mean)	sd
α	2.0	2.037	0.0005
α_c	0.5	0.500	0.0009
β_c	0.3	0.2851	0.004
β	3.0	3.076	0.0524
β_x	1.0	1.043	0.021
α_p	0.5	0.514	0.007
γ	0.2	0.333	0.057
EC	2.5	2.215	0.173
x_{mean}	1	1.152	0.067
ξ_{mean}	0	0.045	0.101
v_{mean}	0	0.357	0.084
σ_x	1	1.004	0.043
σ_ξ	2	1.875	0.081
σ_v	1	1.350	0.083

Table 10: Monte-Carlo Simulation: MCMC Posterior Distributions, BLP model with Moment Conditions

	Mean	Std. Dev.	Min	Max
Facility Characteristics				
Price private-pay	131.81	20.6	93.0	205.0
Registered nurses	3.83	1.3	0.67	9.30
Licensed practical nurses	2.48	1.0	0.12	8.35
Nurse assistants/aids	12.74	3.0	4.20	25.18
Certified medication aides	0.15	0.4	0.0	2.38
Other services weekly hours per beds	12.55	4.1	3.94	34.66
Capacity (number of beds)	123.87	70.3	21	457
Occupancy rate	89.15%	9.7%	36.9%	100%
Ownership				
Government	13.3%	n.a.	n.a.	n.a.
Not-for-profit	33.2%	n.a.	n.a.	n.a.
Cost Data				
Daily Total Labor Cost	1583	1508	65.88	10832
Daily Total Capital Cost	1401	1492	79.67	13803
Hourly Wage Rate	12.72	2.78	5.48	35.0
Rental Rate(%)	6.375	0	6.375	6.375

Table 11: Descriptive Statistics. Number of observed nursing homes = 286

Parameters(MCMC Estimator)	Mean	q0.025	q0.975
β			
Registered nurse hours per bed	-0.094	-0.519	0.342
Licensed practical nurse hours per bed	0.038	-0.443	0.526
Nurse assistant hours per bed	0.145	-0.091	0.385
Other service staff hours per bed	0.031	-0.203	0.262
Government	-0.108	-1.026	0.780
Not-for-profit	0.221	-0.541	0.976
Constant	-7.269	-9.151	-5.861
γ			
Other service staff hours per bed	-0.011	-0.232	0.205
Government	0.088	-0.742	0.913
Not-for-profit	0.156	-0.560	0.843
Constant	-1.395	-2.528	-0.258
σ_l	781.5	745.9	819.6
σ_c	1181	1126	1244
ρ	0.955	0.945	0.963
EC($Annual)	4,015	-2,810	11,023
Parameters of distributions of $\mathbf{x}$			
Registered nurse hours per bed(Log-Normal)			
μ	1.288	1.251	1.329
σ	0.368	0.338	0.407
Licensed practical nurse hours per bed(Log-Normal)			
μ	0.799	0.738	0.862
σ	0.511	0.473	0.553
Nurse assistant hours per bed(Log-Normal)			
μ	2.476	2.419	2.520
σ	0.257	0.230	0.281
Other service staff hours per bed(Log-Normal)			
μ	2.442	2.386	2.483
σ	0.294	0.270	0.321
Government, Not-for-profit, For-profit(Categorical)			
Government	0.129	0.093	0.171
Not-for-profit	0.332	0.280	0.388
For-profit	0.538	0.479	0.595
Parameter(OLS Estimator)	Estimate	SE	
α	3.109	0.0740	
α_c	0.698	1.4880	
β_c	0.662	0.1056	
κ	1.248	0.037	
Mean own price elasticity	4.097		

Iteration: 20,000 times. Burn in: first 10,000 iterations.

Table 12: Estimation Results of the nursing home market

9 Appendix

Appendix A: Identification without cost data

Below, I propose to construct moments that only requires product characteristics for entrant firms. In that case, I cannot condition on $(w_m, r_m, \mathbf{x}_m)$ to identify the choice probability or construct the moment condition for the population that includes both entrant and non-entrant firms. Therefore, the first step of identification only relies on the variation of (w_m, r_m) conditional on $(\mathbf{x}_m, Q_m)$ of entrant firms.

1. Identification of $R_0 \equiv \alpha_{c0}/\left(\alpha_{c0}+\beta_{c0}\right)$.

I first show that the true $\alpha_{c0}/\left(\alpha_{c0}+\beta_{c0}\right)$ is identified. To do so, I consider $(\mathbf{x}_m, Q_m)$ as given. Then, I show that in the model, the combination of w_m, r_m that satisfies

$$\frac{\alpha_{c0}}{\alpha_{c0}+\beta_{c0}} lnw_m + \frac{\beta_{c0}}{\alpha_{c0}+\beta_{c0}} lnr_m = A \tag{93}$$

for a constant A results in the same joint distribution of (p_m, s_m). Given the assumptions and the Equation (93), in Equations (32) and (33), the only remaining variation that determines (p_m, s_m) in the population for all firms that includes entrant and non-entrant firms is the one by (ξ_m, υ_m). This is because all other components of the two equations (32) and (33) are either assumed given or constant in the population. Furthermore, since the the profit in Equation (34), and hence, the entry probability in Equation (35) are only functions of (p_m, s_m), they only depend on the variation of (ξ_m, υ_m). Therefore, variation of (w_m, r_m) satisfying Equation (93) does not change the distribution of (p_m, s_m) of the entrants. On the other hand, if we consider $\alpha_c/\left(\alpha_c+\beta_c\right) \neq \alpha_{c0}/\left(\alpha_{c0}+\beta_{c0}\right)$, then restricting (w_m, r_m) to satisfy

$$\frac{\alpha_c}{\alpha_c+\beta_c} lnw_m + \frac{\beta_c}{\alpha_c+\beta_c} lnr_m = A \tag{94}$$

results in variation of (w_m, r_m) not satisfying Equation (93), thus, (p_m, s_m) of the entrant is not independent to the variation of (w_m, r_m). Therefore, $R_0 \equiv \alpha_{c0}/\left(\alpha_{c0}+\beta_{c0}\right)$ is identified.

2. Identification of the entry probability P_E.

Next, I consider identification of the profit function and the entry probability function P_E. To do so, I consider $\mathbf{x}_m$ as given and choose the combination of (w_m, r_m, Q_m) that satisfies

$$R_0 lnw_m + \left(1-R_0\right) lnr_m + \left(\frac{1}{\alpha_c+\beta_c} - 1\right) ln\left(Q_m\right) = A \tag{95}$$

for a constant A. Then, in Equations (32) and (33), given that the variation of (w_m, r_m, Q_m) satisfy Equation (95), the only remaining exogenous variation comes from (ξ_m, υ_m). As before, because of their independence to $(w_m, r_m, Q_m, \mathbf{x}_m)$, the generated (p_m, s_m) from Equations (32) and (33) in the population, which includes both entrant and nonentrant firms, have the same distribution regardless of the variation in (w_m, r_m, Q_m). Hence, the only variation in the distribution of the profit function comes from the variation of Q_m. Hence, the conditional distribuion of market size Q_m of entrant firms given (p_m, s_m) identifies the entry probability as a function of profit up to the monotone transformation. More formally, let us denote

$$B \equiv (1-\alpha_c-\beta_c)\,/\frac{\alpha_c+\beta_c}{\alpha}, B_0 \equiv (1-\alpha_{c0}-\beta_{c0})\,/\frac{\alpha_{c0}+\beta_{c0}}{\alpha_0}$$

$$C \equiv \frac{\alpha_c+\beta_c}{\alpha}, C_0 \equiv \frac{\alpha_{c0}+\beta_{c0}}{\alpha_0}.$$

Then,

$$P_E\left(Qs\left[(1-\alpha_c-\beta_c)\,p+\frac{\alpha_c+\beta_c}{\alpha}\frac{1}{(1-s)}\right]\right) = P_E\left(CQs\left[Bp+\frac{1}{1-s}\right]\right)$$

Note that in the population that includes both entrants and non-entrants, from Equations (32), (33) and (95),

$$f(s,p,Q|w,r,A,x) = f(s,p|w,r,Q,A,x)\,f(Q|w,r,A,x) = f(s,p|A,x)\,f(Q|w,r,A,x)$$

Taking expectation with respect to (w, r),

$$f(s,p,Q|A,x) = f(s,p|A,x)\,E_{(w,r)}\left[f(Q|w,r,A,x)\right]$$

Then, given the entry probability (add: f_E is the joint distribution of entrants),

$$\frac{f_E(s,p,Q|A,x)}{P_E\left(CQs\left[Bp+\frac{1}{1-s}\right]\right)} = f(s,p|A,x)\,f(Q|A,x) \tag{96}$$

This identifies B_0 and thus, the profit function is identified up to the multiplicative constant $C = (\alpha_c+\beta_c)\,/\alpha$:

$$\pi = CsQ\left[B_0 p+\frac{1}{1-s}\right].$$

Furthermore, by using Equation (96), given any $C \neq 0$, I can identify the function g that satisfy the following property[23]:

$$g\left(CQs\left[B_0p+\frac{1}{1-s}\right]\right)C_E=P_{E0}\left(C_0Qs\left[B_0p+\frac{1}{1-s}\right]\right) \tag{97}$$

Without loss of generality, I set $C=1$. Then, g derived from above satisfies

$$g\left(Qs\left[B_0p+\frac{1}{1-s}\right]\right)C_E=P_{E0}\left(C_0Qs\left[B_0p+\frac{1}{1-s}\right]\right)$$

for some $C_E\neq 0$.

Finally, I can derive C_E given $C=1$ using the below equation:

$$\widehat{E}\left[g\left(sQ\left[B_0p+\frac{1}{1-s}\right]\right)C_E|w,r,Q\right]=\widehat{E}\left[I_E=1|w,r,Q\right]. \tag{98}$$

Then, I can identify the function h that satisfies

$$h\left(p,s,Q\right)=C_Eg\left(sQ\left[B_0p+\frac{1}{1-s}\right]\right)=P_{E0}\left(C_0Qs\left[B_0p+\frac{1}{1-s}\right]\right).$$

I then use the derived function $h\left(\right)$ for the construction of the population moment condition. That is let

$$\begin{aligned}
\varphi_{1m} &\equiv lns_m-ln\left(1-s_m\right)+p_m\alpha-\mathbf{x}_m\boldsymbol{\beta}\\
\varphi_{2m} &\equiv ln\left[p_m-\frac{1}{\left(1-s_m\right)\alpha}\right]-\left[R_0\left(lnw_m-ln\alpha_c\right)+\left(1-R_0\right)\left(lnr_m-ln\beta_c\right)\right]\\
&\quad -\left[\left(\frac{1}{\alpha_c+\beta_c}-1\right)ln\left(Q_ms_m\right)+\mathbf{x}_m\boldsymbol{\gamma}\right].
\end{aligned}$$

Then,

$$\begin{aligned}
E\left[\frac{\varphi_{1m}}{h\left(p_m,s_m,Q_m\right)}\times\mathbf{z}_m\right] &= 0\\
E\left[\frac{\varphi_{2m}}{h\left(p_m,s_m,Q_m\right)}\times\mathbf{z}_m\right] &= 0
\end{aligned}$$

These population moments identify the remaining parameters.

[23] The identification of this section implies that entry probability function g is nonparametrically identified, and thus the entry cost is not identified. However, as modeled in the next section, once we impose parametric assumption on function g, the entry cost is identified.

Appendix B: Identification with cost data

I have already shown that the price coefficient of the logit market share function, α and the true parameters of the Cobb-Douglas cost function, α_c and β_c are identified from Equations (28) and (31) solely from the data on entrant firms. Furthermore, I can recover the true cost C_m^* using Equation (27).

Next, for the entrant firms, I recover the firm specific exogenous component of demand and cost, (δ_m, η_m) where

$$\delta_m \equiv \mathbf{x}_m \boldsymbol{\beta} + \xi_m = lns_m - lns_{0m} + p_m \alpha,$$

and

$$\eta_m \equiv \mathbf{x}_m \boldsymbol{\gamma} + \upsilon_m$$

for the parameterization in Equations (15) and (3). Using Equation (4) and $MR_m = MC_m$, η_m can be recovered as

$$\begin{aligned} \eta_m &= ln \left[q_m \left(p_m - \frac{1}{(1-s_m)\,\alpha} \right) \right] \\ &\quad - \frac{\alpha_c}{\alpha_c + \beta_c} \left(lnw_m - ln\alpha_c \right) - \frac{\beta_c}{\alpha_c + \beta_c} \left(lnr_m - ln\beta_c \right) - \frac{1}{\alpha_c + \beta_c} lnq_m \end{aligned}$$

Furthermore, I can recover the true profit as

$$\pi_m = p_m q_m - C_m^*$$

for the entrant firms.

I first show that $\boldsymbol{\beta}_0$ and $\boldsymbol{\gamma}_0$ are identified. This allows us to decompose the firm specific exogenous components (δ_m, η_m) into the observed $(\mathbf{x}_m)$ and unobserved (ξ_m, υ_m) components. To do so, I use the orthogonality condition between the observed product characteristics and the observed demand and supply shocks. Then, I discuss the identification of the entry cost parameter E_c and the entry probability.

In order to recover the joint distribution of $\mathbf{x}_m$, ξ_m and υ_m, I use the orthogonality assumption that $(\mathbf{w}_m, Q_m, \mathbf{x}_m)$ are independent to ξ_m and υ_m in the population that includes both entrants and non-entrants. That is, the joint distribution of $(\mathbf{x}_m, \xi_m, \upsilon_m)$ conditional on $(\mathbf{w}_m, Q_m)$ is specified as follows:

$$f\left(\mathbf{x}_m, \xi_m, \upsilon_m | \mathbf{w}_m, Q_m\right) = f_{\mathbf{x}}\left(\mathbf{x}_m | \mathbf{w}_m, Q_m\right) f_{(\xi,\upsilon)}\left(\xi_m, \upsilon_m\right)$$

where $f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)$ is the density function of the distribution of $\mathbf{x}$ conditional on $(\mathbf{w},Q)$, and $f_{(\xi,\upsilon)}$ is the joint distribution of the demand and cost shocks (ξ,υ).

Note that if all potential entrants enter, and thus their variables are observable, then, (x_m,δ_m,η_m) of all firms are observable. Furthermore, I have $\xi_m=\delta_m-\mathbf{x}_m\boldsymbol{\beta}$, and $\upsilon_m=\eta_m-\mathbf{x}_m\boldsymbol{\gamma}$. If I assume that $\mathbf{x}$ and (ξ,υ) are independent to each other, by applying the change of variables, I obtain

$$f(\mathbf{x},\delta,\eta|\mathbf{w},Q)\,d\mathbf{x}d\delta d\eta=f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)\,f_{(\xi,\upsilon)}(\xi,\upsilon)\,d\mathbf{x}d\xi d\upsilon.$$

Therefore, if all firms are entrants, I can identify $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ from the orthogonality condition

$$f(\mathbf{x},\delta,\eta|\mathbf{w},Q)=f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)\,f_{(\xi,\upsilon)}(\delta-\mathbf{x}\boldsymbol{\beta},\eta-\mathbf{x}\boldsymbol{\gamma}).$$

To do so, I assume that the tail of the joint distribution $f_{(\xi,\upsilon)}(,)$ converges to zero. That is, $f_{(\xi,\upsilon)}(\xi,\upsilon)$ can be made arbitrarily small by making $\xi^2+\upsilon^2$ sufficiently large.

Let $(\mathbf{x},\delta_0,\eta_0)$, $(\mathbf{x}',\delta_0',\eta_0')$ such that $\delta_0=\mathbf{x}\boldsymbol{\beta}_0+\xi$, $\eta_0=\mathbf{x}\boldsymbol{\gamma}_0+\upsilon$, $\delta_0'=\mathbf{x}'\boldsymbol{\beta}_0+\xi$, $\eta_0'=\mathbf{x}'\boldsymbol{\gamma}_0+\upsilon$. Then, $\delta_0-\mathbf{x}\boldsymbol{\beta}_0=\delta_0'-\mathbf{x}'\boldsymbol{\beta}_0=\xi$, and $\eta_0-\mathbf{x}\boldsymbol{\gamma}_0=\eta_0'-\mathbf{x}'\boldsymbol{\gamma}_0=\upsilon$. Then,

$$\begin{aligned}\frac{f(\mathbf{x},\delta_0,\eta_0|\mathbf{w},Q)}{f(\mathbf{x}',\delta_0',\eta_0'|\mathbf{w},Q)} &= \frac{f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)}{f_{\mathbf{x}}(\mathbf{x}'|\mathbf{w},Q)}\frac{f_{(\xi,\upsilon)}(\delta_0-\mathbf{x}\boldsymbol{\beta}_0,\eta_0-\mathbf{x}\boldsymbol{\gamma}_0)}{f_{(\xi,\upsilon)}(\delta_0'-\mathbf{x}'\boldsymbol{\beta}_0,\eta_0'-\mathbf{x}'\boldsymbol{\gamma}_0)}.\\ &= \frac{f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)}{f_{\mathbf{x}}(\mathbf{x}'|\mathbf{w},Q)}\frac{f_{(\xi,\upsilon)}(\xi,\upsilon)}{f_{(\xi,\upsilon)}(\xi,\upsilon)}=\frac{f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)}{f_{\mathbf{x}}(\mathbf{x}'|\mathbf{w},Q)}\end{aligned}$$

On the other hand, for $(\boldsymbol{\beta},\boldsymbol{\gamma})\neq(\boldsymbol{\beta}_0,\boldsymbol{\gamma}_0)$, let $\delta=\mathbf{x}\boldsymbol{\beta}+\xi=\delta_0+\mathbf{x}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)$, $\eta=\mathbf{x}\boldsymbol{\gamma}+\upsilon=\eta_0+\mathbf{x}(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0)$, $\delta'=\mathbf{x}'\boldsymbol{\beta}+\xi=\delta_0'+\mathbf{x}'(\boldsymbol{\beta}-\boldsymbol{\beta}_0)$, $\eta'=\mathbf{x}'\boldsymbol{\gamma}+\upsilon=\eta_0'+\mathbf{x}'(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0)$. Then, $\delta-\mathbf{x}\boldsymbol{\beta}=\delta'-\mathbf{x}'\boldsymbol{\beta}=\xi$, and $\eta-\mathbf{x}\boldsymbol{\gamma}=\eta'-\mathbf{x}'\boldsymbol{\gamma}=\upsilon$, but $\delta-\mathbf{x}\boldsymbol{\beta}_0=\xi+\mathbf{x}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)\neq\delta'-\mathbf{x}'\boldsymbol{\beta}_0=\xi+\mathbf{x}'(\boldsymbol{\beta}-\boldsymbol{\beta}_0)$, and $\eta-\mathbf{x}\boldsymbol{\gamma}_0=\upsilon+\mathbf{x}(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0)\neq\eta'-\mathbf{x}'\boldsymbol{\gamma}_0=\upsilon+\mathbf{x}'(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0)$. Therefore, for $(\boldsymbol{\beta},\boldsymbol{\gamma})\neq(\boldsymbol{\beta}_0,\boldsymbol{\gamma}_0)$,

$$\begin{aligned}\frac{f(\mathbf{x},\delta,\eta|\mathbf{w},Q)}{f(\mathbf{x}',\delta',\eta'|\mathbf{w},Q)} &= \frac{f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)}{f_{\mathbf{x}}(\mathbf{x}'|\mathbf{w},Q)}\frac{f_{(\xi,\upsilon)}(\delta-\mathbf{x}\boldsymbol{\beta},\eta-\mathbf{x}\boldsymbol{\gamma})}{f_{(\xi,\upsilon)}(\delta'-\mathbf{x}'\boldsymbol{\beta},\eta'-\mathbf{x}'\boldsymbol{\gamma})}.\\ &= \frac{f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)}{f_{\mathbf{x}}(\mathbf{x}'|\mathbf{w},Q)}\frac{f_{(\xi,\upsilon)}(\xi+\mathbf{x}(\boldsymbol{\beta}-\boldsymbol{\beta}_0),\upsilon+\mathbf{x}(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0))}{f_{(\xi,\upsilon)}(\xi+\mathbf{x}'(\boldsymbol{\beta}-\boldsymbol{\beta}_0),\upsilon+\mathbf{x}'(\boldsymbol{\gamma}-\boldsymbol{\gamma}_0))}\neq\frac{f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)}{f_{\mathbf{x}}(\mathbf{x}'|\mathbf{w},Q)}\end{aligned}$$

for $\mathbf{x}'=a\mathbf{x}$ for sufficiently large scalar $a>0$. This is due to the assumption of the tails of the joint distribution converging to zero. Therefore, if I set (δ,η) and (δ',η') as above, then

$$\frac{f(\mathbf{x},\delta,\eta|\mathbf{w},Q)}{f(\mathbf{x}',\delta',\eta'|\mathbf{w},Q)}=\frac{f_{\mathbf{x}}(\mathbf{x}|\mathbf{w},Q)}{f_{\mathbf{x}}(\mathbf{x}'|\mathbf{w},Q)}.$$

for any $(\mathbf{x}, \mathbf{x}', \xi, \upsilon)$ if and only if $(\boldsymbol{\beta}, \boldsymbol{\gamma}) = (\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)$.

Next, I consider the case where only a subset of firms enter. Then, as before, I specify the entry probability to be a function of the profit minus the entry cost E_C. That is, the entry cost is specified as $P_E(\pi - E_C, \mathbf{w}, Q) = P_E(pq - C^* - E_C, \mathbf{w}, Q)$. Furthermore, let I_E be the dummy for the entrant firm. That is, $I_E = 1$ if the firm is an entrant and $I_E = 0$ if the firm is a non-entrant. Then,

$$\begin{aligned} & f(\mathbf{x}, \delta, \eta, I_E|\mathbf{w}, Q) \\ = & \Big[P_E(pq - C^* - E_C, \mathbf{w}, Q)\, I_E + \Big(1 - P_E(pq - C^* - E_C, \mathbf{w}, Q)\Big)(1 - I_E)\Big] \\ & \times f_{\mathbf{x}}(\mathbf{x}|\mathbf{w}, Q)\, f_{(\xi,\upsilon)}(\xi, \upsilon), \end{aligned} \tag{99}$$

where $(\xi, \upsilon) = (\delta - \mathbf{x}\boldsymbol{\beta}, \eta - \mathbf{x}\boldsymbol{\gamma})$. Recall that for entrant firms, $(\mathbf{w}, Q, \mathbf{x}, \delta, \eta, p, q, C^*)$ are either observable in the data or can be recovered from the procedure explained in Section 3.

Given the restriction of the model, profit can be expressed as $\pi(w, Q, \delta, \eta)$. Furthermore, the population density function conditional on $(\mathbf{w}, Q)$ for entrants can be expressed as

$$f(\mathbf{x}, \delta, \eta, 1|\mathbf{w}, Q) \equiv f(\mathbf{x}, \delta, \eta, I_E|\mathbf{w}, Q, I_E = 1).$$

I also specify the entry probability for the potential entrant in market m, P_{Em} to be an increasing function of profit, i.e.

$$P_{Em} = P_E(\pi_m, \mathbf{w}_m, Q_m), \ \frac{\partial g}{\partial \pi_m} > 0.$$

Hence, for entrants, $I_{em} = 1$, and from Equation (99),

$$\frac{f(\mathbf{x}_m, \delta_{0m}, \eta_{0m}, 1|\mathbf{w}_m, Q_m)}{P_E(\pi(\mathbf{w}_m, Q_m, \delta_{0m}, \eta_{0m}), \mathbf{w}_m, Q_m)} = f_{\mathbf{x}}(\mathbf{x}_m|\mathbf{w}_m, Q_m)\, f_{(\xi,\upsilon)}(\xi_m, \upsilon_m), \ m \in \mathcal{E}. \tag{100}$$

Finally, given that I can obtain entry probability given $(\mathbf{w}_m, Q_m)$ in the data, I have the following equation with $\widehat{P}(\mathbf{w}_m, Q_m)$ being the probabilty of a market having one monopolist given the exogenous and observable market condition $(\mathbf{w}_m, Q_m)$:

$$E[g(\pi_m, \mathbf{w}_m, Q_m)|\mathbf{w}_m, Q_m] = \widehat{P}(\mathbf{w}_m, Q_m) \tag{101}$$

Then, I define the identification of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ as follows: the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are identified if the Equation

$$\frac{f\left(\mathbf{x}_m, \mathbf{x}_m\boldsymbol{\beta} + \xi_m, \mathbf{x}_m\boldsymbol{\gamma} + \upsilon_m, 1|\mathbf{w}_m, Q_m\right)}{g\left(\pi\left(\mathbf{w}_m, Q_m, \mathbf{x}_m\boldsymbol{\beta} + \xi_m, \mathbf{x}_m\boldsymbol{\gamma} + \upsilon_m\right), \mathbf{w}_m, Q_m\right)} = f_{\mathbf{x}}\left(\mathbf{x}_m|\mathbf{w}_m, Q_m\right) f_{(\xi,\upsilon)}\left(\xi_m, \upsilon_m\right), \ m \in \mathcal{E}. \tag{102}$$

and Equation (101) hold for any entrant firm in the population if and only if $(\boldsymbol{\beta}, \boldsymbol{\gamma}) = (\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)$. From now on until the end of the proof, the variables without subscript are the ones of the entrant firms. That is, throughout the proof, I only use data on entrant firms for identification.

$$f\left(\mathbf{x}, \delta_0, \eta_0, 1|\mathbf{w}, Q\right) = g\left(\pi\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}_m, Q_m\right) f_{\mathbf{x}}\left(\mathbf{x}|\mathbf{w}, Q\right) f_{(\xi,\upsilon)}\left(\delta_0 - \mathbf{x}\boldsymbol{\beta}, \eta_0 - \mathbf{x}\boldsymbol{\gamma}\right)$$

$$f\left(\mathbf{x}, \delta_0, \eta_0, 1|\mathbf{w}, Q\right) = g_0\left(\pi\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}_m, Q_m\right) f_{\mathbf{x}0}\left(\mathbf{x}|\mathbf{w}, Q\right) f_{(\xi,\upsilon)0}\left(\delta_0 - \mathbf{x}\boldsymbol{\beta}_0, \eta_0 - \mathbf{x}\boldsymbol{\gamma}_0\right) \tag{103}$$

Hence,

$$\frac{g_0\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}_m, Q_m\right)}{g\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}_m, Q_m\right)} = \frac{f_{\mathbf{x}}\left(\mathbf{x}|\mathbf{w}, Q\right)}{f_{\mathbf{x}0}\left(\mathbf{x}|\mathbf{w}, Q\right)} \times \frac{f_{(\xi,\upsilon)}\left(\delta_0 - \mathbf{x}\boldsymbol{\beta}, \eta_0 - \mathbf{x}\boldsymbol{\gamma}\right)}{f_{(\xi,\upsilon)0}\left(\delta_0 - \mathbf{x}\boldsymbol{\beta}_0, \eta_0 - \mathbf{x}\boldsymbol{\gamma}_0\right)} \tag{104}$$

Then, if $(\delta_0 - \mathbf{x}\boldsymbol{\beta}_0, \eta_0 - \mathbf{x}\boldsymbol{\gamma}_0)$ is independent of $\mathbf{x}$, then $(\delta_0 - \mathbf{x}\boldsymbol{\beta}, \eta_0 - \mathbf{x}\boldsymbol{\gamma})$ is independent of $\mathbf{x}$ if and only if $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ and $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0$. Therefore, $(\boldsymbol{\beta}_0, \boldsymbol{\gamma}_0)$ is identified. Therefore, $f_{(\xi,\upsilon)} = f_{(\xi,\upsilon)0}$. Therefore,

$$\frac{g_0\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}, Q\right)}{g\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}, Q\right)} = \frac{f_{\mathbf{x}}\left(\mathbf{x}|\mathbf{w}, Q\right)}{f_{\mathbf{x}0}\left(\mathbf{x}|\mathbf{w}, Q\right)}$$

In addition, since the LHS is not a function of $\mathbf{x}$,

$$\frac{f_{\mathbf{x}}\left(\mathbf{x}|\mathbf{w}, Q\right)}{f_{\mathbf{x}0}\left(\mathbf{x}|\mathbf{w}, Q\right)} = \varphi\left(\mathbf{w}, Q\right)$$

Then,

$$g_0\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}_m, Q_m\right) = \frac{g\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}_m, Q_m\right)}{\varphi\left(\mathbf{w}, Q\right)}$$

That is, the entry probability and the conditional distribution of $\mathbf{x}$ are unidentified. For additional identification, I observe the entry probability conditional on $(\mathbf{w}, Q)$. Therefore, by setting

$$\frac{E\left[g\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}, Q\right)|\mathbf{w}, Q\right]}{\varphi\left(\mathbf{w}, Q\right)} = E\left[g_0\left(\pi_0\left(\mathbf{w}, Q, \delta_0, \eta_0\right), \mathbf{w}, Q\right)|\mathbf{w}, Q\right]$$

, I identify $\varphi\left(\mathbf{w}, Q\right)$, and thus, g_0. Imposing a parametric assumption on g identifies E_c.