INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

541
2020

The Equivalence Between Sequential and Simultaneous Firm Decisions

Francisco Silva y Samir Mamadehussene.

www.economia.puc.cl • ISSN (edición impresa) 0716-7334 • ISSN (edición electrónica) 0717-7593

The Equivalence Between Sequential and Simultaneous Firm Decisions

Samir Mamadehussene* Francisco Silva†

Abstract

When firms compete by choosing two strategic variables (e.g. quality and price), the timing under which firms make their decisions (simultaneous vs sequential choice of the strategic variables) plays a critical role, as the equilibrium may be drastically different depending on the timing that is assumed. We rely on the marketing and psychology literatures that provide well-established evidence that consumers do not consider all products in a market, i.e. consumers form "consideration sets". Under this assumption, we find that in markets where (i) firms' strategies do not influence the consideration set formation, and (ii) firms are sufficiently uncertain regarding the rivals that each consumer considers, the equilibrium of the game in which firms choose the strategic variables sequentially is close to the equilibrium of the simultaneous game. Moreover, the equilibrium of the simultaneous game does not depend on whether or not consumers consider all available alternatives. Therefore, we argue that the analysis of these markets can be performed with standard models provided that the simultaneous timing is used (even if firms make their decisions sequentially).

*Católica-Lisbon School of Business and Economics, 1649-023 Lisboa, Portugal, samir@ucp.pt
†Instituto de Economia, Pontificia Universidad Catolica de Chile, Chile, franciscosilva@uc.cl

1 Introduction

In markets where firms compete by choosing two strategic variables, such as advertising and price (e.g. Butters 1977), the quality of the product and its price (e.g. Shaked and Sutton 1982), capacity and price (e.g. Deneckere and Peck 1995), a price policy and a price (e.g. Thisse and Vives 1988), or the price and the way that the price information is framed (e.g. Piccione and Spiegler 2012), the timing under which firms make their decisions plays a crucial role. The approach taken by the literature is that if one of the two variables appears to be a longer-run decision variable, the sequential timing is used, i.e., first, firms simultaneously choose the longer-run decision variable, and then, after observing each others' choices, simultaneously choose the other variable. For example, when the two variables are price and quality, quality seems like a longer-run decision, so that most of the literature (although there are some exceptions) assumes a sequential timing. If none of the variables stands out as being a long-run variable, a simultaneous timing is more appropriate, where all firms choose both strategic variables at the same time.

The literature is not always in agreement regarding which timing to use. Take the example of the literature on informative advertising, where firms are able to send advertisements that include the price of their products. While Butters [1977], Stegeman [1991], and Robert and Stahl [1993] model the advertising and pricing decisions as simultaneous, arguing that advertisements are messages that firms send that include the price, Ireland [1993] and McAfee [1994] use a sequential timing, in which firms choose advertising first and prices second, because, as is stated in Ireland [1993], "the quantity of information is decided in the first period as a level of advertising and then the second period involves the insertion of price into the advertisement".

There are several other instances where the literature is not consensual regarding the timing of the decision variables. For example, in the literature of obfuscation, in which firms choose prices and the complexity of the price information, Carlin [2009] and Ellison and Wolitzky [2012] consider a simultaneous timing, whereas Wilson [2010] and Gu and Wenzel [2014] consider a sequential timing, where firms choose obfuscation first and prices second. Chioveanu and Zhou [2013] and Mamadehussene [2020a] argue that both timings are reasonable.

This lack of consensus is particularly unsettling given that the equilibria of the various models may be drastically different, depending on the timing that is assumed. For example, the results from Carlin [2009] are not robust to the sequential timing, while the results from Wilson [2010] are not robust to the simultaneous timing.[1] Our main contribution is to propose an argument in favor of the simultaneous timing that does not rely on determining whether one of the two decision variables is more longer-run than the other.

Most of the industrial organization literature makes the implicit assumption that consumers are aware of all available products. However, this assumption has been refuted by a vast body of evidence, which we document below, that seems to support the idea that each consumer only considers a subset of the available products. We find that i) when we assume that consumers have "consideration sets", the equilibrium of the game in which firms choose the strategic variables sequentially is "close" to the equilibrium of the simultaneous game; and ii) the equilibrium of the simultaneous game does not depend on whether or not consumers consider all available alternatives. By combining i) and ii), we argue that the analysis of markets under the more realistic assumption that consumers do not consider all available products can be performed with standard models (that assume that consumers do consider all products), provided that the simultaneous timing is used (even if firms actually make their decisions sequentially).

The marketing and psychology literatures provide well established evidence that consumers do not consider all products in a given market before making a purchase (Hauser and Wernerfelt 1990; Roberts and Lattin 1991). There are many reasons for consumers not

[1]We discuss Carlin [2009] in some detail in section 4, while the non-robustness to the timing of the model of Wilson [2010] is discussed by Wilson himself.

to consider all available alternatives. First, they might not be aware of all the products in a given market (e.g. Laurent et al. 1995). For example, Honka et al. [2017] focus on the 18 largest financial institutions in the United States and find that consumers are, on average, aware of less than 7 of those banks. Similarly, Tsai and Honka [2018] focus on the top 22 brands in the U.S. auto insurance industry and find that consumers are, on average, aware of only 12 of those brands. Second, even when consumers are aware of all firms that operate in a market, they may just fail to remember them (Biehal and Chakravarti 1986; Lynch Jr et al. 1988; Nedungadi 1990). Finally, the fact that it is costly to evaluate the various alternatives may lead consumers to choose to consider only a subset of the available products (Simon 1955; Shugan 1980).

In light of this evidence, the economics literature has recently taken into account the fact that consumers do not consider all available alternatives. For example, Goeree [2008] analyses the U.S. personal computer market and argues that consumers are unlikely to be aware of all PCs for sale. She finds that traditional demand models, that assume that consumers are aware of all brands in the market, overstate price elasticities. The decision theory literature also analyzes the implications of the fact that consumers do not consider all available alternatives (Masatlioglu et al. 2012; Manzini and Mariotti 2014; Lleras et al. 2017).

Hart [1985] was perhaps the first to allow for consumers to consider only a subset of the available alternatives in a model of competition among firms. In his model, each consumer is assumed to be interested in only a subset of the potential brands available. He argues that "a possible justification for the assumption that each consumer chooses among a small number of brands is that, due to information costs, the consumer is unaware of the prices and/or qualities of the majority of the brands available". He finds that assuming that consumers are not aware of all brands is fundamental to obtain monopolistic competition.

Given that relaxing the assumption that consumers are aware of all firms plays such an important role in monopolistic competition, understanding the implications of such assumption in oligopolistic markets is fundamental. Perloff and Salop [1985] discuss the implications of consumers not being aware of all firms in a model of oligopolistic price competition. In their model, the equilibrium price, denoted by $p(n)$, is a function of the number of firms in the market, n. They analyze the implications of consumers being imperfectly informed about which firms operate in the market, and provide clear predictions: "suppose each consumer is aware of only $k < n$ brands. Then there are $m = \binom{n}{k}$ equal-sized submarkets, each consisting of k brands [...]. Given [...] symmetry assumptions [...], each submarket is identical and equilibrium is achieved at the k-firm equilibrium price $p(k)$". In light of this result, the assumption that consumers are aware of all firms that operate in the market seems innocuous, as it only affects the definition of the parameter n, which can be interpreted not as the total number of firms that operate in the market, but rather as the number of firms in the consumer's consideration set. By contrast, we find that the assumption that consumers are aware of all firms is not innocuous when firms compete by choosing two strategic variables, as opposed to the one decision variable (price) model of Perloff and Salop [1985].

Finally, because our results provide a rationale for using the simultaneous timing, they also help validate the results of some of the simultaneous models in the literature. It is often the case that models where firms choose the strategic variables simultaneously are more tractable than their sequential counterpart (e.g. Thisse and Vives 1988). Moreover, the analysis of sequential games is usually limited to duopolies, in order to restrict the number of subgames of the second stage (e.g. Champsaur and Rochet 1989). For these reasons, sometimes the simultaneous timing is used, even though it could be argued that firms choose the strategic variables sequentially. Indeed, the simultaneous timing is used by Deneckere and Peck [1995] (to model the choice of capacity and price), Vandenbosch and Weinberg [1995] (to model the choice of two dimensions of product quality), Wiggins and Lane [1983] and Eliaz and Spiegler [2011] (to model the choice of product quality and advertising), Dubovik and Janssen [2012] and Bachi and Spiegler [2018] (to model the choice of quality and price) and Janssen and Parakhonyak [2013] and Mamadehussene [2019] (to model the choice of price and whether to provide a price-matching guarantee). Naturally, the concern is that the results obtained in these papers may fail to hold if firms choose the strategic variables in sequence. Our results mitigate this concern; we argue that, provided the consumers' consideration sets do not include all alternatives available in the market, the results from the aforementioned papers are valid even if firms choose the strategic variables sequentially.

The rest of the paper proceeds as follows. In section 2 we provide an illustrative example, in which firms choose pure strategies in the first stage. Section 3 provides our main results. Section 4 discusses an application in which firms play mixed strategies. Finally, in section 5 we discuss the conditions that are required for our results to hold, and we provide examples of market environments that satisfy such conditions.

2 An example with pure strategies in the first stage

We start with a simple example to illustrate two key points: i) in standard models, where consumers are assumed to consider all available alternatives, the equilibrium may be drastically different depending on whether it is assumed that firms make decisions simultaneously or sequentially; ii) in a setting where consumers' consideration set is a subset of the available firms, the equilibrium is not very sensitive to the choice of timing.

Consider a market where N firms compete to supply a homogeneous product. They face the same marginal cost, which we normalize to zero. Each firm decides whether to carry the product in its store ($\alpha = 1$) or to have the product only available to order ($\alpha = 0$). Besides choosing product availability, firms also set prices.

There is a unit mass of consumers, each demanding, at most, one unit of the product if their valuation (normalized to one) is not exceeded. Each consumer has a consideration set of 2 firms (this example considers a duopoly setting as it is standard when analyzing sequential games). For simplicity, we assume that the mass of consumers with consideration set (a, b) is $\frac{1}{\binom{N}{2}}$ for any firms a, b. At the end of this section we show that the results are robust when we depart from this assumption.

There are two types of consumers. A fraction $\lambda < \frac{1}{2}$ of consumers have urgency to get the product. If the product is available at only one store (out of the 2 stores in their consideration set), they purchase the product at that store. If the product is available in both stores, they purchase at a random store. Finally, if the product is not available at any store, they do not purchase.

The remaining $(1 - \lambda)$ mass of consumers order the product from the store in their consideration set that has the lowest price.

2.1 When consumers consider all firms, i.e., N=2

Simultaneous choice of availability and price

We start by analyzing the case in which firms choose availability and prices simultaneously. The profit of a firm when it chooses availability α and price p, given that the other store chooses availability $\tilde{\alpha}$ and price $\tilde{p}$ is:

$$\pi(\alpha, p/\tilde{\alpha}, \tilde{p}) = \alpha(1 - \tilde{\alpha})\lambda p + \alpha\tilde{\alpha}\frac{\lambda}{2}p + \mathbb{1}[p < \tilde{p}](1 - \lambda)p + \mathbb{1}[p = \tilde{p}]\frac{1 - \lambda}{2}p$$

It is straightforward that the profit is increasing in α. This result is intuitive: when prices and availability are chosen at the same time, a firm has nothing to gain by making the product unavailable in its store, because availability is costless. In equilibrium both firms carry the product ($\alpha = 1$). The game then becomes similar to Varian (1980): a fraction λ of consumers buy at a random store, whereas the remaining consumers buy at the lowest priced store. Each firm chooses a price from price distribution F_{Sim}, characterized below and earns profit $\frac{\lambda}{2}$.

$$F_{Sim}(p) = \frac{2-\lambda}{2(1-\lambda)} - \frac{\lambda}{2(1-\lambda)}\frac{1}{p} \quad \text{for } p \in \left[\frac{\lambda}{2-\lambda}, 1\right] \tag{1}$$

Sequential choice of availability and price

We now analyze the equilibrium under the sequential timing: in the first stage, firms simultaneously choose availability; in the second stage, after firms' availability choices are observed, firms simultaneously set prices.

There are four subgames: (1) both firms carry the product ($\alpha_1 = 1, \alpha_2 = 1$); (2) neither firm carries the product ($\alpha_1 = 0, \alpha_2 = 0$); (3) only firm 1 carries the product ($\alpha_1 = 1, \alpha_2 = 0$); (4) only firm 2 carries the product ($\alpha_1 = 0, \alpha_2 = 1$). Subgames (3) and (4) are symmetric. We examine each of the first 3 subgames below.

Both firms carry the product

This subgame is similar to Varian (1980): a fraction λ of consumers buys at a random store, whereas the remaining consumers buy at the lowest priced store. Each firm chooses a price from price distribution F_{Sim} characterized in (1) and each firm earns profit $\frac{\lambda}{2}$.

Neither firm carries the product

This subgame is a Bertrand game: consumers with urgency to get the product do not purchase, whereas the remaining consumers purchase from the firm that sets the lowest price. In equilibrium, both firms charge marginal cost and make zero profit.

Only firm 1 carries the product

In this subgame, all consumers with urgency to get the product purchase from firm 1. The remaining consumers purchase from the firm with the lowest price. The equilibrium price distributions of firms 1 and 2, respectively denoted by F_1 and F_2, are characterized below.

$$F_1(p) = 1 - \frac{\lambda}{p} \quad \text{for } p \in [\lambda, 1] \tag{2}$$

$$F_2(p) = \frac{1}{1-\lambda} - \frac{\lambda}{1-\lambda}\frac{1}{p} \quad \text{for } p \in [\lambda, 1] \tag{3}$$

In equilibrium, firm 1 makes profit λ, whereas firm 2 makes profit $\lambda(1-\lambda)$.

We can now characterize the equilibrium in the first stage. We focus on equilibria under which firms play pure strategies in the first stage. In equilibrium, only one firm carries the product. There are two asymmetric equilibria in the first stage: $(\alpha_1 = 1, \alpha_2 = 0)$ and $(\alpha_1 = 0, \alpha_2 = 1)$. Without loss of generality, we focus on the first one.

Comparison

Under the sequential timing, the availability strategy of a firm influences the pricing decision of its rival. When firm 2 does not carry the product, firm 1 is assured to sell to all consumers with urgency to get the product. This provides incentives for firm 1 to charge higher prices, which benefits both firms.

In contrast, under the simultaneous timing the availability strategy of a firm does not influence the pricing strategy of its rival. Therefore, in the equilibrium of the simultaneous game both firms carry the product.

Figure 1 illustrates this point. In the sequential game, the fact that firm 2 does not carry the product softens price competition: both firms charge higher prices than the prices charged in the simultaneous equilibrium.

The equilibrium predictions of this game are very different depending on the timing under which firms choose their strategies. Under the simultaneous timing, both firms carry the product, whereas under the sequential timing only one firm does so. Moreover, both firms

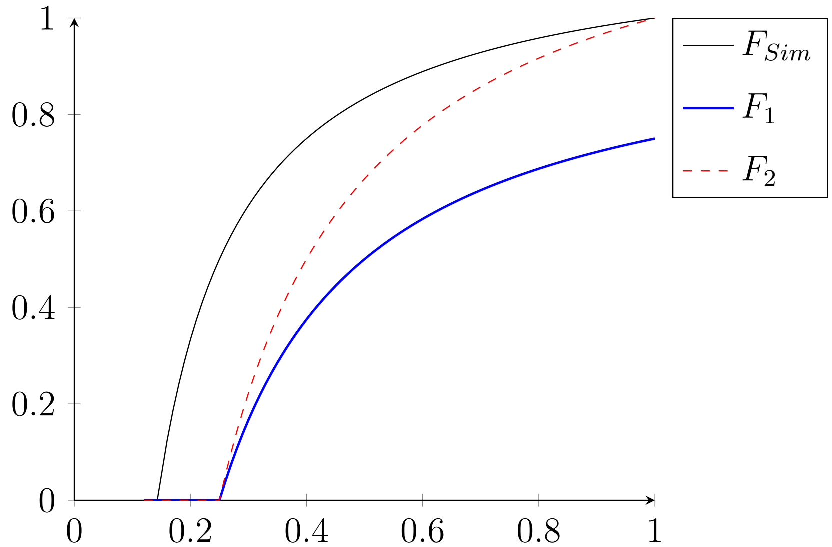


Figure 1: Equilibrium price distributions for $\lambda = 0.25$

make higher profits under the sequential timing.

2.2 When consumers' consideration set is a subset of all firms, i.e., $N > 2$

Simultaneous choice of availability and price

The equilibrium of the simultaneous game is the same as in the case in which $N = 2$. When firms are choosing their strategies, they take into account that, for each consumer, they will compete against one rival. The fact that firms face uncertainty regarding the identity of their rival does not change their incentives. Any symmetric equilibrium when $N = 2$ is also a symmetric equilibrium for $N > 2$ (in the next session, we show that this is true in general). Therefore, in equilibrium all firms carry the product in their stores ($\alpha = 1$), and each firm chooses a price from price distribution F_{Sim} characterized in (1).

Sequential choice of availability and price

We focus on equilibria under which firms play pure strategies in the first stage, as we did for the case $N = 2$. A subgame is characterized by the number of firms that choose not to carry the product in their stores, which we denote by r (i.e. r is the number of firms that choose $\alpha = 0$). The complete analysis of all subgames is provided in the appendix; in the text, we only summarize the main results. Let $r^*(N)$ denote the equilibrium number of firm that choose $\alpha = 0$, when there are N firms in the market.

Lemma 1 *For any $N > 2$: $r^*(N) \geq r^*(N+1)$*

As the number of firms in the market increases, fewer firms choose $\alpha = 0$. Hence, an increase in N leads to a significant reduction in the proportion of firms that do not carry the product: not only does the denominator (N) increase, the numerator (r) also decreases. If N is not too small, then in equilibrium all firms choose to carry the product in their stores.

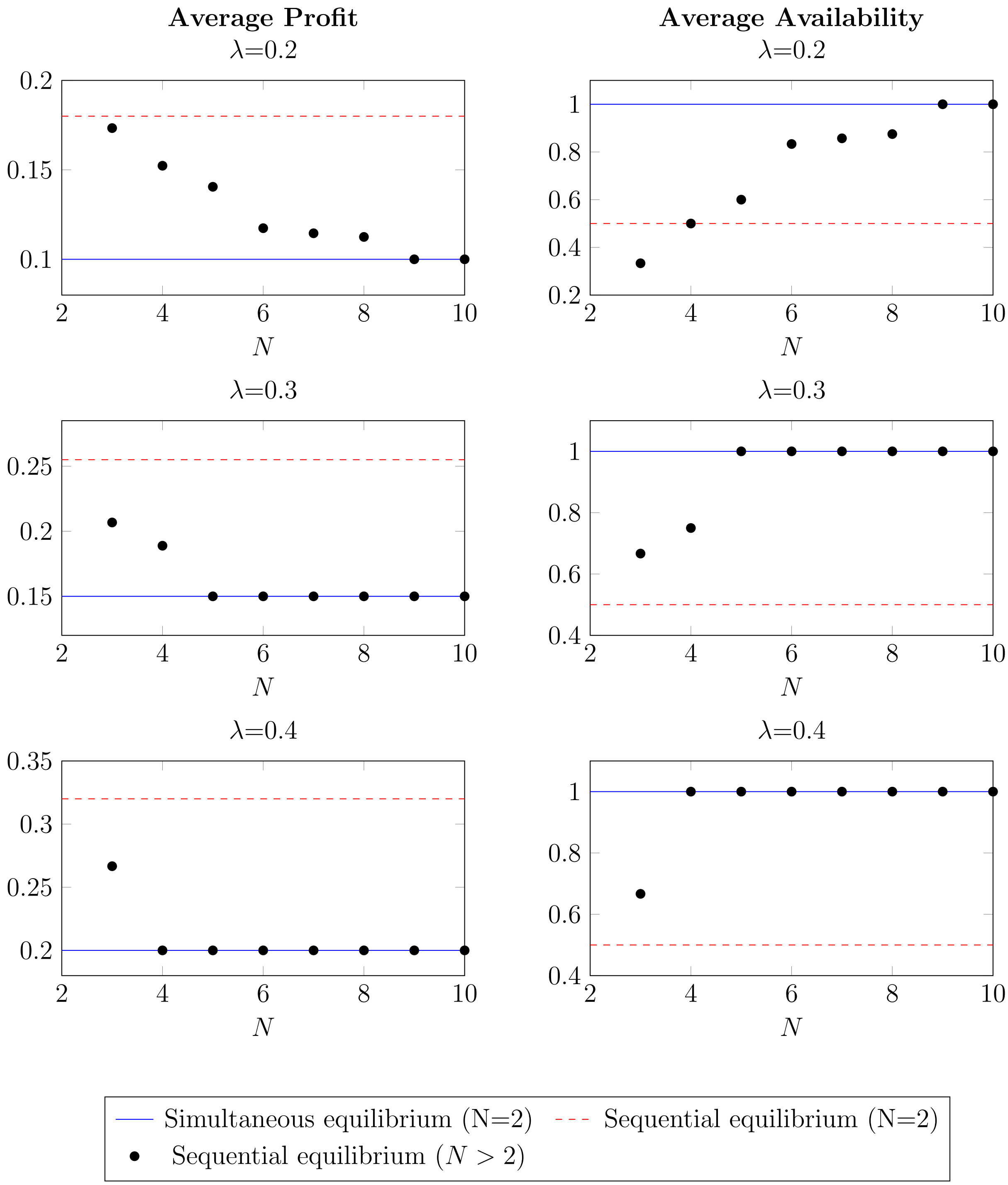
Proposition 1 *If $N > \frac{2(1-\lambda)}{\lambda}$ then $r^*(N) = 0$*

Comparison

When $N > 2$, there are no fundamental differences in the model predictions that arise from changing the timing. In both the sequential and the simultaneous timing, most firms choose to carry the product in their stores. The fact that consumers only consider a subset of all firms that operate in the market makes timing less important - when the number of firms in the market, N, is not too small, the equilibrium in the simultaneous and sequential games coincide.[2] This point is illustrated in figure 2: when $N > 2$, the equilibrium in the sequential game is close to the equilibrium in the simultaneous game.

To see why that is, let us momentarily go back to the case $N = 2$. When availability and price are chosen simultaneously, a firm's choice of α does not influence the rival's pricing strategy. Therefore, firms have no incentives to choose $\alpha = 0$.

[2] We show this result in general in the next section, when there is a continuum of firms.



Average Profit
Average Availability
λ=0.2
λ=0.3
λ=0.4
N
Simultaneous equilibrium (N=2)
Sequential equilibrium (N=2)
Sequential equilibrium (N > 2)

Figure 2

This is not true when availability decisions are made before firms set prices. Let us fix the availability strategy of firm 1 to $\alpha_1 = 1$. If firm 2 also chooses $\alpha_2 = 1$, then in the second stage each firm sets prices taking into account that there is a 0.5 probability that it sells to a consumer with urgency to get the product. If, instead, firm 2 chooses $\alpha_2 = 0$, then in the second stage firm 1 will set prices taking into account that it will sell to all consumers with urgency to get the product. Therefore, when firm 2 chooses $\alpha_2 = 0$ the mass of consumers that go directly to firm 1 doubles, which provides large incentives for firm 1 to charge higher prices. When firm 2 chooses not to carry the product, it faces the downside of not selling to consumers with urgency to get the product, but it is able to extract a larger surplus from the remaining consumers, given that firm 1 is charging very high prices.

Let us now return to the case $N > 2$ and fix the availability strategies of the first $N-1$ firms to $\alpha_1 = \alpha_2 = ... = \alpha_{N-1} = 1$. If firm N also chooses $\alpha_N = 1$ then in the second stage each firm i sets prices taking into account that there is a 0.5 probability that it sells to a consumer who has firm i in its consideration set and faces urgency to get the product. If, instead, firm N chooses $\alpha_N = 0$, then in the second stage each firm $i \in \{1, ..., N-1\}$ will set prices taking into account that, in case firm i is lucky enough that its rival is firm N, an event with probability $\frac{1}{N-1}$, it will sell to all consumers with urgency to get the product. If, however, the rival is not firm N, there is a 0.5 probability that firm i sells to a consumer with urgency to get the product. Hence, when $\alpha_N = 0$ there is a $\frac{1}{N-1} + \frac{N-2}{N-1}\frac{1}{2} = \frac{1}{2}\frac{N}{N-1}$ probability that firm i sells to a consumer who has firm i in its consideration set and faces urgency to get the product. Therefore, the probability that a consumers who has firm i in its consideration set and faces urgency to get the product buys from firm i rises from $\frac{1}{2}$ (when $\alpha_N = 1$) to $\frac{1}{2}\frac{N}{N-1}$ (when $\alpha_N = 0$), an increase of $\frac{1}{2(N-1)}$. The jump in probability becomes insignificant as N increases. Therefore, when $N > 2$, choosing $\alpha = 0$ will have a very small effect on the rivals' pricing strategies.

When consumers consider only a subset of all firms that operate in the market, even when availability and prices are set sequentially, the incentives that firms face when choosing their availability strategies are similar to the incentives faced in the simultaneous game: choosing $\alpha = 0$ will have (almost) no impact on the rivals' pricing strategies. This point is illustrated in Figure 3. The dashed red line depicts the price distribution that firms play in case all firms choose $\alpha = 1$ (i.e. when they fall in the subgame $r = 0$); the solid black line depicts the price distribution that firms with $\alpha = 1$ play in case exactly one firm chooses $\alpha = 0$ (i.e. when they fall in the subgame $r = 1$). When $N = 2$, there is a large discrepancy between the prices that a firm with $\alpha = 1$ charges when $r = 0$ vs when $r = 1$: indeed, when $r = 1$ the price distribution even has a mass point at the consumer reservation value. In contrast, when $N > 2$ there is no longer a mass point at such price and having one firm choosing $\alpha = 0$ leads the remaining firms to charge prices only slightly higher than they charge when all firms play $\alpha = 1$.

2.3 Nonuniform consideration sets

The key argument behind our results is that when consumers consider only a subset of the firms that compete in the market, the first-stage decision of a firm will have a moderate impact on the second-stage strategies of its rivals. For simplicity, we have considered the symmetric case, in which all consideration sets have the same measure. This implies that the first-stage decision of any firm will have the same impact on all its rivals.

Our main argument is not driven by this simplifying assumption. The important part is that firms must be sufficiently uncertain regarding which rivals they are competing with. This point is better illustrated with an example. Suppose there are two stores located in the North side of the city (N_1 and N_2) and two stores located in the South side (S_1 and S_2). Each consumer has two stores in his consideration set: the store near his home and the store near his workplace. A proportion $\frac{1}{4}$ of consumers lives close to each store. However, home and work locations are correlated. In particular, a consumer who lives near a store in the North (e.g. N_1) is more likely to work near the other store in the North (N_2) than near a store in the South. More concretely, a consumer who lives near N_1 works near N_2 with

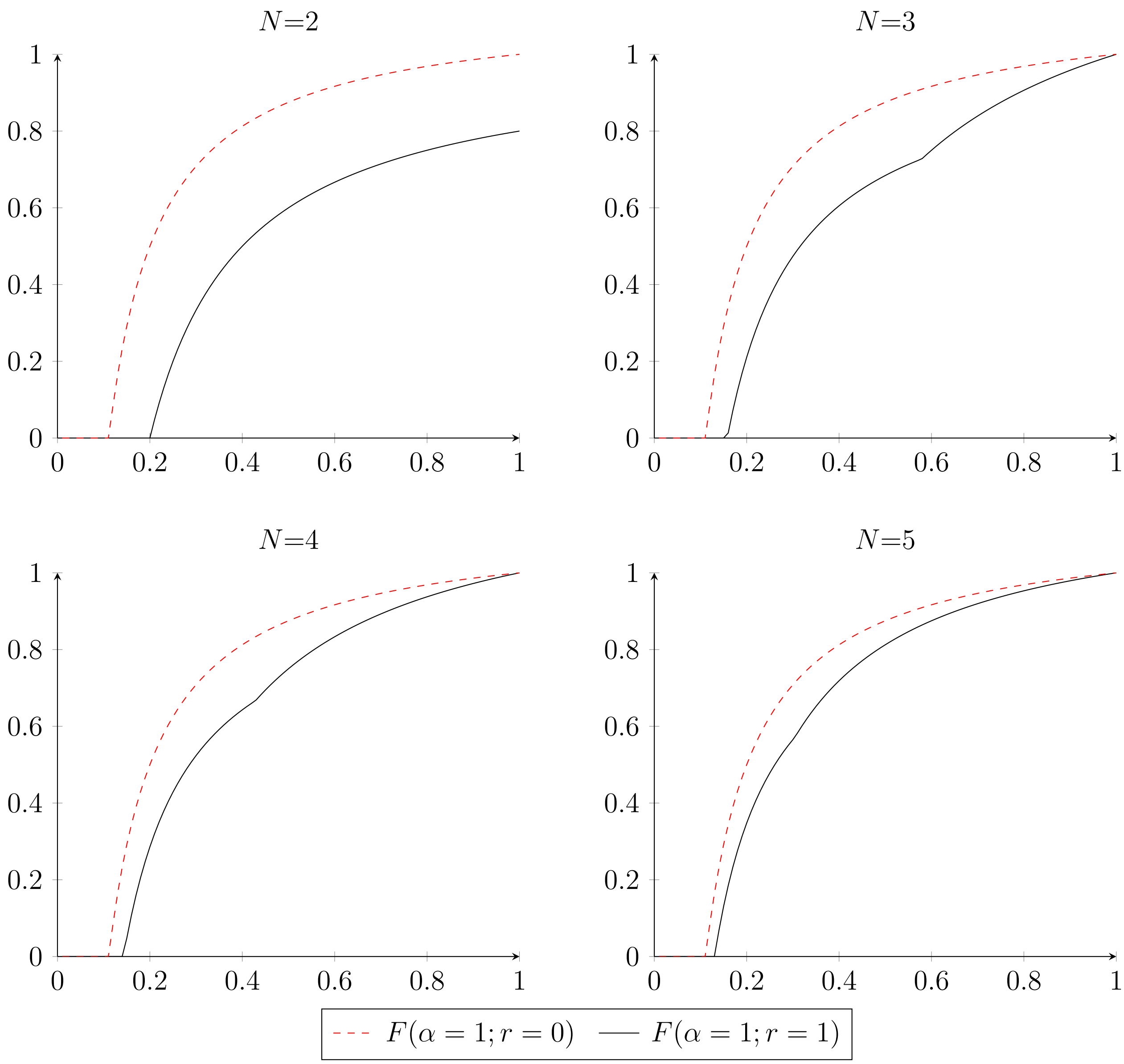


Figure 3: Price distributions for $\lambda = 0.2$

probability $\left(\frac{1}{3}+\epsilon\right)$ and works near S_1 (and S_2) with probability $\left(\frac{1}{3}-\frac{\epsilon}{2}\right)$.

Notice that when $\epsilon = \frac{2}{3}$ the North and South are two independent markets: stores in the North never compete with stores in the South. Clearly, our results would not apply to such case. If $\epsilon = 0$ we have the uniform consideration set, in which case we have shown that our results apply. In this section we show that the results are robust to nonuniform consideration sets ($\epsilon > 0$).

Let us consider $\lambda > 0.4$, in which case under the uniform consideration set assumption ($\epsilon = 0$) the equilibrium of the sequential game coincides with the equilibrium of the simultaneous game: all firms play $\alpha = 1$. In this setting ($\epsilon = 0$), suppose a firm plays $\alpha = 0$ in the first stage (while the remaining firms play $\alpha = 1$). In the pricing stage, a firm with $\alpha = 1$ will choose prices taking into account that there is a $\frac{1}{3}$ chance that its rival has $\alpha = 0$ and a $\frac{2}{3}$ chance that its rival has $\alpha = 1$. This will lead to a small impact in their pricing strategies (compared to the case in which all firms play $\alpha = 1$). For this reason, no firm finds it optimal to play $\alpha = 0$.

Proposition 2 *Let $\lambda > 0.4$. If $\epsilon < 0.04$ then in the equilibrium of the sequential game all firms play $\alpha = 1$.*

The intuition for why the equilibrium does not change when ϵ is small is better understood with an example. Let $\epsilon = 0.01(6)$, so that out of all consumers with N_1 in their consideration set, 35% also have store N_2 in their consideration set, whereas 32.5% have S_1 (S_2) in their consideration set. Suppose firm N_1 plays $\alpha = 0$ in the first stage (whereas the remaining firms play $\alpha = 1$). Firm N_2 would benefit more from the choice of N_1 than would firms S_1 and S_2. When firm N_2 is setting prices, it takes into account that there is a 35% chance that its rival has $\alpha = 0$ and a 65% chance that its rival has $\alpha = 1$. When firm N_1 sets $\alpha = 0$ it impacts the pricing incentives of N_2 in a very similar way than in the uniform consideration set case ($\epsilon = 0$). Although the choice of $\alpha = 0$ by N_1 will have a larger impact on the pricing strategy of N_2 than on the pricing strategies of S_1 and S_2, the impact is still small for all firms' pricing strategies. Therefore, each firm finds is optimal to play $\alpha = 1$.

When ϵ is large, the uncertainty regarding the identity of the rival is small (for example, if ϵ is close to $\frac{2}{3}$, firm N_1 knows that a consumer who visits its store will most likely also visit N_2). In this case, the first stage decision of N_1 will have a large impact on the pricing strategy of N_2. Therefore, while our results hold even when there is correlation among firms in the likelihood of appearing in a consumer's consideration set, that correlation must not be too large.

3 Model and results

We compare two models where N firms compete. In either model, there are two stages; in the first stage, each firm chooses some two-variable policy $z_i = (x_i, y_i) \in X \times Y \equiv Z \subseteq \mathbb{R}^J$ for some $J \geq 2$, while in the second stage, consumers take some action as a function of the observed policy vector $z \in Z^N$. What is different between the models is the timing of the first stage; in the *simultaneous* model, each firm i chooses (x_i, y_i) simultaneously, while in the *sequential* model, each firm i first chooses x_i and only then chooses y_i after observing vector $x = (x_1, ..., x_N)$.

The payoff of policy $z_i \in Z$ for firm i depends on the policies other firms choose. In particular, we assume that, for each vector $z_{-i} \in Z^{N-1}$, the payoff of policy $z_i \in Z$ is given by

$$\pi^i (z_i, z_{-i}) = b^i (z_i, z_{-i}) - c^i (z_i),$$

where $c^i (z_i) \in \mathbb{R}$ represents the cost of policy z_i, while $b^i (z_i, z_{-i})$ represents the benefit. Function b^i has a particular structure that depends on the consumers' behavior. In particular, we assume that in the second stage of each game, there is a continuum of consumers and each of them considers only a set γ of $n \leq N$ firms and their policies (we call this the consumer's consideration set). If firm i does not belong to the consideration set of the consumer, the benefit it gets from that consumer is 0; if it does, the benefit is given by $v^i(z_i, z_{\gamma-i})$, where $z_{\gamma-i} \in Z^{n-1}$ represents the vector of all policies considered by the consumer with consideration set γ other than z_i. Let Γ denote the set of all subsets of the N firms with size n, i.e., the set of all possible consideration sets. The measure of any given consideration set $\gamma \in \Gamma$ is denoted by $\mu(\gamma) \in [0,1]$. While each firm knows μ, it does not know the consideration set of each consumer.

When analyzing the simultaneous game, the literature typically considers the symmetric case (for example, Carlin [2009], which we discuss in the following section). Firms are symmetric if functions v_i and c_i are independent of i and if there is $k \in (0,1)$ such that $\mu(\gamma) = k$ for all $\gamma \in \Gamma$.

Proposition 3 *Assume that firms are symmetric and that the measure of consumers per firm is equal to one,*[3] *and let $n \geq 2$. If a symmetric strategy profile is a Nash equilibrium of the simultaneous game when $N = n$, it is also a Nash equilibrium for any $N > n$.*

Proposition 3 generalizes the result from the previous section in stating that the total number of firms N does not influence symmetric equilibria when the game is simultaneous; what is relevant is the number of firms n each consumer considers. By contrast, in general, the total number of firms N does influence the set of symmetric subgame perfect equilibria of the sequential game as can be seen by the example of section 4. However, as we show below, the two sets of equilibria converge when the number of players is large. In particular, we find that even when firms are heterogeneous, the set of Nash equilibrium outcomes of the simultaneous game is the same as the set of subgame perfect equilibrium outcomes of the sequential game if there is a continuum of agents.[4]

[3]Assuming that the measure of consumers increases at the same rate as N keeps the competitive pressures among firms constant with the total number of firms. This assumption is standard (e.g. Burdett and Judd 1983).

[4]The argument adapts Fudenberg and Levine [1988] to our setting.

3.1 Continuum of firms

There is a continuum of firms $i \in [0,1]$ and each chooses a policy $z_i \in Z$. Once policies are chosen, each consumer is matched with n firms, just like before. In this setting, $\Gamma = [0,1]^n$ and μ represents the continuous distribution over Γ. Firms i and j are symmetric if $v^i = v^j$, $c^i = c^j$ and $\mu(i, \beta) = \mu(j, \beta)$ for all $\beta \in [0,1]^{n-1}$ that does not include neither i nor j.

We allow firms to be heterogeneous: there is a mapping $t : [0,1] \to T$, where $t(i)$ represents firm i's (public) type. Firms with the same type are symmetric. We slightly abuse notation and rewrite functions v^i, c^i and π^i as $v^{t(i)}$, $c^{t(i)}$ and $\pi^{t(i)}$. Without loss of generality, we assume that firms with the same type follow the same strategy.[5] We also assume that T is finite, so that there are many firms that behave in the same way. Therefore, we can write

$$\pi^t(z_i, f) = E_\omega\left(v^t(z_i, \omega) \,|f\right) - c^t(z_i),$$

where $\omega = (\omega_1, ..., \omega_{n-1})$ represents the vector of policies that each type t-firm is competing against and $f : T \to \Delta Z$ is such that each $f^t \in \Delta Z$ represents the distribution of actions over Z that is induced by the type t-firms' play.

In the simultaneous game, a Nash equilibrium is a function $\sigma : T \to \Delta Z$ such that, for all $t \in T$ and $\sigma^t \in \Delta Z$,

$$\int\limits_{z_i \in Z} \sigma^t(z_i)\, \pi^t(z_i, \sigma)\, dz_i \geq \pi^t(z_i', \sigma)$$

for all $z_i' \in Z$.

In the sequential game, a strategy for each type t-firm is a pair $\left(\sigma_x^t, \sigma_y^t\right)$, where $\sigma_x^t \in \Delta X$,

[5]Notice that this does not preclude the possibility of asymetric equilibria. For example, if one wants to allow a set of completely homogeneous firms to follow two different strategies for example, one can simply divide the set into two distinct sets (so T would have two elements instead of just one even though the firms are homogeneous).

while $\sigma_y^t(x_i, f_x) \in \Delta Y$, where $f_x : T \to \Delta X$ and each $f_x^t \in \Delta X$ is the observed distribution of actions played over set X by firms of type t. In words, in a SPE, the second period choice of any type t-firm depends only on its initial choice of $x_i \in supp(f_x^i)$ and on the empirical distributions generated by all the other types. Let $g_y(f_x, \sigma_y) : T \to \Delta Y$ be such that each $g_y^t(f_x, \sigma_y) \in \Delta Y$ denotes the distribution over Y generated by t-type firms, when firms play $\sigma_y : T \to \Delta Y$ in the second stage after having observed $f_x : T \to \Delta X$ in the first stage. Formally, for each $y_i \in Y$,

$$g_y^t(f_x, \sigma_y)(y_i) = \int\limits_{x_i \in X} f_x^t(x_i)\, \sigma_y^t(x_i, f_x)(y_i)\, dx_i.$$

Finally, (σ_x, σ_y) is a SPE if i) for all $t \in T$ and for all $f_x : T \to \Delta X$ and $x_i \in supp(f_x^t)$,

$$\int\limits_{y_i \in Y} \sigma_y^t(x_i, f_x)(y_i)\, \pi^t((x_i, y_i), (f_x, g_y(f_x, \sigma_y)))\, dy_i \geq \pi^t((x_i, y_i'), (f_x, g_y(f_x, \sigma_y)))$$

for all $y_i' \in Y_i$ and ii)

$$\int\limits_{x_i \in X} \int\limits_{y_i \in Y} \sigma_x^t(x_i)\, \sigma_y^t(x_i, \sigma_x)(y_i)\, \pi^t((x_i, y_i), (\sigma_x, g_y(\sigma_x, \sigma_y)))\, dy_i dx_i$$
$$\geq \int\limits_{y_i \in Y} \sigma_y^t(x_i', \sigma_x)(y_i)\, \pi^t((x_i', y_i), (\sigma_x, g_y(\sigma_x, \sigma_y)))\, dy_i$$

for all $x_i' \in X_i$. We assume that for any $f_x : T \to (\Delta X)$, there is always some $\sigma_y(\cdot, f_x)$ that satisfies condition i), i.e., we assume that there is always some second stage Nash equilibrium for any first period play.

The object $\sigma : T \to \Delta Z$ is a SPE outcome if there is a SPE (σ_x, σ_y) such that

$$\sigma^t((x_i, y_i)) = \sigma_x^t(x_i)\, g_y^t(\sigma_x, \sigma_y)(y_i)$$

for all $t \in T$ and for all $(x_i, y_i) \in Z$.

Proposition 4 *Every strategy profile $\sigma : T \to \Delta Z$ is a Nash equilibrium of the simultaneous game if and only if it is a SPE outcome of the sequential game.*

The argument is slightly more involved than the example of the previous section because it also applies to mixed strategies. Essentially, we rely on the fact that, when there is a continuum of agents, the distribution of play that is induced by the players' strategy profile is equal to the empirical distribution that is observed. We illustrate in the next section by revisiting Carlin [2009] and analyzing it under the assumption that consumers only consider a subset of all the firms in the market.

4 Revisiting Carlin [2009]

Carlin [2009] studies how a homogeneous good market is affected when firms can choose different price complexities. We revisit Carlin [2009] because it is an example of a model whose mixed strategy equilibria under the different timing assumptions is very different when consumers consider every firm. By contrast, we show that the equilibria are very similar when consumers only consider a subset of the firms. Below, we provide a brief overview of Carlin's model.

There are N homogeneous firms and a continuum of consumers of measure 1. Each consumer is willing to pay up to $v > 0$ to purchase the good, which can be produced by each firm at zero cost. In the first stage of the game, each firm i simultaneously chooses a price $p_i \in [0, v]$ and a complexity value $k_i \in \left[\underline{k}, \overline{k}\right]$ where $\underline{k} < \overline{k}$. In the second stage, the consumer chooses from which firm to buy. Firms' price complexity influences the probability that the consumer is informed; in particular, that probability is given by $\mu\left(k_1, ..., k_N\right)$, where μ is decreasing with each k_i. Informed consumers know every firms' price and buy from the lowest one. However, uninformed consumers, who do not know which firm has the lowest price, are equally likely to purchase from any firm. Under the assumption that $\frac{\partial^2 \mu}{\partial k_i \partial k_j} = 0$ for any two firms i and j, Carlin finds that there is a symmetric Nash equilibrium of the game played by the firms where each firm plays a mixed strategy. In particular, each firm i chooses $k_i = \overline{k}$ with some probability α and chooses $k_i = \underline{k}$ with probability $(1-\alpha)$; if $k_i = \overline{k}$, firm i randomizes in its price choice over some interval $[\overline{p}, v]$, while if $k_i = \underline{k}$, firm i randomizes in its price choice over some interval $\left[\underline{p}, \overline{p}\right]$, where $\underline{p} < \overline{p} < v$ (proposition 1 in Carlin 2009).[6]

Carlin's main result is, in his own words, that "the model predicts that as competitive pressures rise in an industry [...] firms will respond by adding more complexity to their prices". This result is summarized in his Proposition 2 that shows that, as the number of firms increases, the proportion α of firms who choose the maximum complexity level increases.

4.1 Sequential choice of price and complexity

Consider an alternative version of Carlin's model where the N firms first select the price complexity and only then select prices. Consider some arbitrary subgame, where agents have played $(k_1, ..., k_N)$ in the first stage. The subgame is similar to Varian [1980]: uninformed consumers buy at a random firm, whereas informed consumers buy at the firm with the lowest price. In the symmetric Nash equilibria of the subgame, it follows that each firm randomizes in its price decision, and the profit of each firm is given by

$$\frac{1-\mu\left(k_1, ..., k_N\right)}{n} v.$$

Because the profit is increasing with k_i for all i, the only symmetric subgame perfect equilibrium is one where each firm i chooses the maximum complexity level in the first stage $(k_i = \overline{k})$ and then randomizes over prices in the second stage in interval $\left[\underline{\varphi}, v\right]$, where $\underline{\varphi}$ is

[6]The statement of proposition 1 in Carlin [2009] is stated in a slightly different way; this way of presenting the results is more convenient for our comparison.

such that

$$\frac{1-\mu\left(\overline{k},...,\overline{k}\right)}{n}v=\left(\mu\left(\overline{k},...,\overline{k}\right)+\frac{1-\mu\left(\overline{k},...,\overline{k}\right)}{n}\right)\underline{\varphi}.$$

Notice that the main result from Carlin [2009] vanishes when the sequential timing is used. It is no longer true that as competitive pressures rise firms respond by making their prices more complex; indeed, under the sequential timing, the level of complexity is always maximum and independent of the total number of competitors.

4.2 The model with consideration sets

We now study the properties of our model in this context, where each firm has $n-1$ competitors even though there are a total of $N>n$ firms in the market. The game is as follows: in the first stage, each of the N firm chooses its price complexity level; in the second stage, after observing what each of the N firms has chosen, each firm chooses its price. The consideration set of each consumer consists of $n<N$ firms, and the measure of each feasible consideration set is the same, i.e., there is a measure $\frac{1}{\binom{N}{n}}$ of consumers with any given combination of n firms. The probability that each consumer is informed only depends on the complexity levels of the firms in his consideration set. If the consumer is informed, he purchases from the firm with the lowest price, among those in his consideration set. If the consumer is uninformed, he purchases from a random firm among those in his consideration set.

For simplicity, and to make the comparison between this model and Carlin's model (where $n=N$) clearer, we assume that $k_i\in\left\{\underline{k},\overline{k}\right\}$ for all i. We also let $\widehat{\mu}\left(w\right)$ denote the probability that the consumer is informed, given that w of the n considered firms chose the largest complexity level (while $n-w$ firms chose the smallest one). For exposition, we refer to firms which have chosen the largest (smallest) complexity level as "high type" ("low type") firms.

4.2.1 Symmetric equilibria

Let us consider the second stage subgames. Notice that, unlike the case where $n = N$, the second stage subgames are not symmetric, because, if selected, a high type firm will face a larger probability that the consumer is uninformed than a low type firm. As a result, to focus on symmetric Nash equilibria in this context is to focus on Nash equilibria where, for each subgame, all high type firms follow one strategy and all low type firms follow another strategy. The second stage symmetric Nash equilibrium depends on the number of high type firms, which we denote by r.

Using standard arguments (e.g. Varian 1980), we have the following results. If there are $r = 0$ high type firms, then low type firms randomize on their price choice over interval $[p_0, v]$; if there are $r = N$ high type firms, then high type firms randomize over interval $[p_N, v]$; if the number of high type firms r is such that $0 < r < N - 1$, then high type firms randomize over interval $[p_r, v]$, while low type firms randomize over interval $[\widehat{p}_r, p_r]$; finally, if $r = N - 1$, then high type firms randomize over interval $[p_{N-1}, v]$, while the low type firm randomizes over interval $[p_{N-1}, \overline{p}_r]$, where $\overline{p}_r < v$.

To study how firms choose the complexity level of the first stage, it is necessary to identify p_r for all $r = 0, 1, ..., N$. When $r = 0$ or $r = N$, its calculation is straightforward by using the indifference of low and high types respectively: p_0 is such that

$$\frac{1 - \widehat{\mu}(0)}{n} v = \left(\widehat{\mu}(0) + \frac{1 - \widehat{\mu}(0)}{n} \right) p_0,$$

while p_N is such that

$$\frac{1 - \widehat{\mu}(N)}{n} v = \left(\widehat{\mu}(N) + \frac{1 - \widehat{\mu}(N)}{n} \right) p_N.$$

When there are both high and low types, it is slightly more complicated as each firm faces uncertainty over which type of competitors she faces if considered by the consumer. Let

$g\left(\widehat{w},\widehat{r},\widehat{n},\widehat{N}\right)$ denote the probability that exactly $\widehat{w}$ high type firms are chosen to be a part of a group of $\widehat{n}$ firms, when there is a total of $\widehat{N}$ firms and $\widehat{r}$ of them are high type firms. Consider any subgame where there is a total of $r \geq 1$ high type firms. When a high type firm chooses a price of v, its expected profit is given by $\overline{X}(r)\,v$, where

$$\overline{X}(r) = \frac{n}{N}\sum_{w=0}^{n-1} g(w, r-1, n-1, N-1)\,\frac{1-\widehat{\mu}(w+1)}{n}.$$

When it chooses a price of p_r, its expected profit is given by $\overline{Y}(r)\,p_r$, where

$$\overline{Y}(r) = \begin{cases} \overline{X}(r) + \frac{n}{N} g(n-1, r-1, n-1, N-1)\,\widehat{\mu}(n) & \text{if } 0 < r < N-1 \\ \overline{X}(N-1) + \frac{n}{N}\sum_{w=0}^{n-1} g(w, N-2, n-1, N-1)\,\widehat{\mu}(w+1) & \text{if } r = N-1 \end{cases}.$$

Therefore, it follows that $p_r = \frac{\overline{X}(r)}{\overline{Y}(r)} v$ for all $0 < r < N$. Notice also that the expected profit of a low type firm at any subgame where there are a total of $r < N$ high type firms is given by $\underline{Y}(r)\,p_r$, where

$$\underline{Y}(r) = \begin{cases} \frac{n}{N}\left(\widehat{\mu}(0) + \frac{1-\widehat{\mu}(0)}{n}\right) & \text{if } r = 0 \\ \frac{n}{N}\left(\begin{array}{c} \sum_{w=0}^{n-1} g(w, r, n-1, N-1)\,\frac{1-\widehat{\mu}(w)}{n} + \\ g(n-1, r, n-1, N-1)\,\widehat{\mu}(n-1) \end{array}\right) & \text{if } 0 < r < N-1 \\ \frac{n}{N}\left(\widehat{\mu}(n-1) + \frac{1-\widehat{\mu}(n-1)}{n}\right) & \text{if } r = N-1 \end{cases}.$$

In the first period, each firm randomizes between selecting the largest and the lowest complexity level. Let $\alpha \in [0,1]$ be the probability that each firm i selects $k_i = \overline{k}$. It follows that if $\alpha \in (0,1)$, then

$$\sum_{r=0}^{N-1}\binom{N-1}{r}\alpha^r(1-\alpha)^{N-1-r}\,\overline{X}(r+1)\,v = \sum_{r=0}^{N-1}\binom{N-1}{r}\alpha^r(1-\alpha)^{N-1-r}\,\underline{Y}(r)\,p_r, \quad (4)$$

where the left hand side (LHS) represents the expected payoff if $k_i = \overline{k}$, while the right hand side (RHS) represents the expected payoff if $k_i = \underline{k}$.

4.2.2 Comparison to Carlin's original model

Unlike what happens if $n = N$ where all firms choose maximum complexity ($\alpha = 1$), in general when $n < N$ we have that $\alpha \in (0, 1)$, so that, ex-post, some firms choose high levels of complexity while some others choose low levels of complexity, just like in Carlin's paper. As α is typically interior, one can also study how an increase in the number of competitors influences it; recall that Carlin's main result is that an increase in the number of competitors (n) increases the probability that each firm chooses the largest complexity level.[7]

Equation (4) characterizes α. A general analytic characterization of how α changes with n is prohibitively complex. We use the following function: $\widehat{\mu}(w) = \gamma_0 - \gamma_1 \frac{w}{n}$ for $\gamma_0 \in (0, 1)$ and $\gamma_1 \in (0, \gamma_0)$ (notice that this function satisfies the assumptions from Carlin's model). We show numerically that α is increasing in n. Figure 4 presents some numerical examples to illustrate this result (the examples in the figures have fixed $N = 20$). It is also straightforward to show that, for any n, when $N \to \infty$ the equilibrium α converges to the α from Carlin's model, corroborating the results of the previous section.

The reason why the timing matters less and less when N grows is as follows. With the simultaneous timing and mixed equilibria, each firm is best responding against lotteries that are played by the other firms. If $n = N$ but the timing becomes sequential, firms best respond to the actual observed play when they get to the second period. What generates the discrepancy in the results when $n = N$ is precisely that the observed play is a specific realization of the first period strategy profile, i.e., each firm faces no uncertainty in the second period about the first period choices of their competitors. That changes when $N > n$. In that case, there is still uncertainty in the second period even after observing the first period

[7] Carlin's main result is that as competitive pressures increase, so does the probability that firms choose complex prices. Competitive pressures are measured by the size of the consumer consideration set, n.

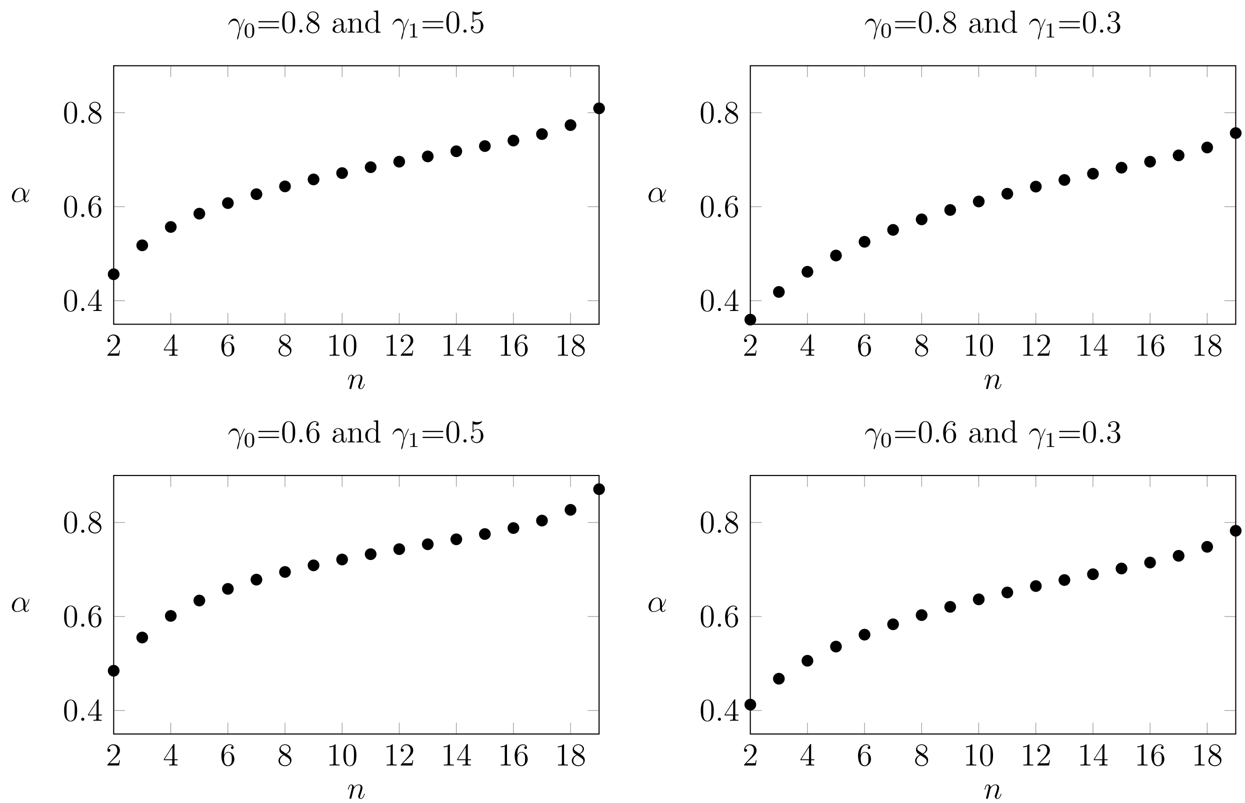


Figure 4

choices of all firms, because each firm does not know which firms are its competitors. Therefore, each firm is best responding to a lottery of possible competitors' actions. Essentially, the uncertainty generated by not knowing the identity of the competitors is strategically equivalent to not knowing which pure strategy is actually played by each firm.

5 Discussion

We complete this article by discussing the conditions that are required for our results to hold, and providing examples of market environments that satisfy such conditions.

5.1 Conditions that are necessary for our results

1. Industries with sufficiently many firms

Because our results require that consumers consider a strict subset of the firms in the market, they do not hold in environments characterized by a very small number of firms, such as the carbonated soft drink market which was, for a long time, a near duopoly between Coca-Cola and Pepsi (Fosfuri and Giarratana 2009). Instead, our findings apply in industries where there are sufficiently many firms. For example, it has been documented that consumers are not aware of all firms that operate in the financial industry (Honka et al. 2017), the auto insurance industry (Tsai and Honka 2018) or the ground coffee industry (Draganska and Klapper 2011).

2. Firms' first-stage decision does not influence consideration sets

We have assumed that the consideration set of a consumer is fixed even before firms choose their strategies. This is an important assumption, as the equilibrium may be very different if firms' first-stage move can influence consumers' consideration sets. Our results would not apply if one wants to model strategies that are likely to influence the consideration set formation, such as advertising or the choice of location. Instead, our results apply to modeling firms' strategies that are not perceived to influence the consideration set formation. For example, if one wants to model firms' choice of quality and price, it is usually assumed that consumers need to search a firm in order to learn the quality of its products (e.g. Anderson and Renault 1999).

3. Firms are sufficiently uncertain regarding the rivals that are competing for each consumer

While our results hold even when there is correlation in the likelihood that firms appear in the consideration set (as we discussed in section 2.3), such correlation cannot be too large. For example, suppose firms A and B always appear together in the consideration set (i.e. either both firms are in the consideration set or neither is). In this case, these firms are always competing with each other, and therefore the first-stage strategy of A will have a large impact on the second-stage strategy of B. Naturally, our results do not hold in this setting. Similarly, if one major firm is much more likely to appear in the consideration set than the remaining firms, the first-stage decision of such major firm will have a strong impact on the strategies of the remaining firms. Instead, our results apply when firms' are sufficiently uncertain regarding the rivals that are competing for each consumer.

At first glance, it may seem that this condition is not satisfied when consumers need to travel to visit a store. Indeed, if two stores are located next to each other, they may always appear together in consideration sets. However, this is not necessarily the case. Houde [2012] finds that in the market for gasoline consumers consider stores along their commuting paths. Mamadehussene [2020b] reports a similar shopping behavior in the market for tires. When consumers' consideration sets include stores along their commuting paths, each store faces uncertainty regarding the rivals that are competing for each consumer. As an illustrative example, consider the roads depicted in Figure 5. Although stores A and B are located within a small distance, firm A is not necessarily competing with firm B. Indeed, depending on the consumer's commuting path, a consumer who drives by store A may also drive by C, D or E.

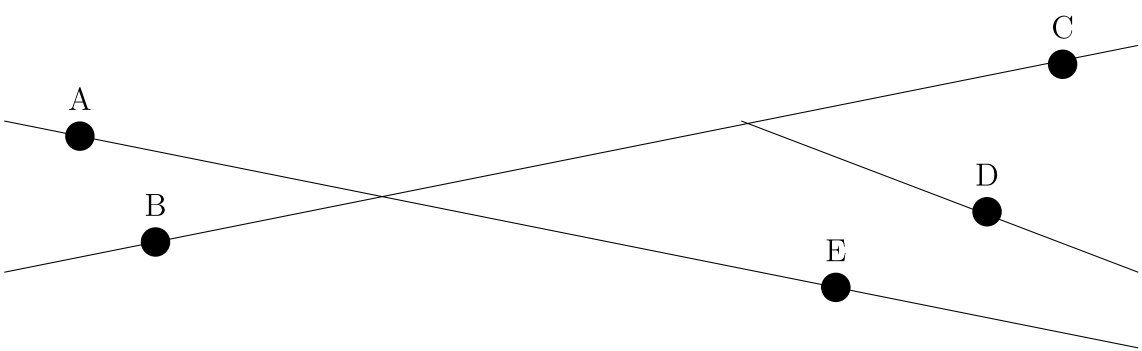


Figure 5: Each line represents a road

5.2 Examples of environments that satisfy the above conditions

One example where all three conditions seem to be satisfied is the competition among chain stores. Condition 1) is likely to be satisfied because there are many different chains competing in the market. Regarding condition 2), there are many variables that these firms choose that do not seem to influence the consideration set formation. Take the example of the fast food industry. The consideration set of each consumer will likely include the restaurants located nearby; in this case, the consideration set is unaffected by firms' choices such as product line or capacity.

Because chains typically follow national pricing policies[8] (i.e., they choose the same policy in every store regardless of where the store is located), condition 3) is also likely to be satisfied. To see this more clearly, let us go back to the example of the fast food industry, and consider firms' decisions regarding product lines and prices. There are 20 fast food chains that each operate more than 2,000 locations in the U.S. The consideration set of each consumer plausibly includes only the stores located nearby. Whereas some consumers have both Subway and McDonald's in their consideration set, others have Subway and Wendy's. Because Subway chooses the same product line for all its stores, when making such decision it faces uncertainty regarding the identity of the rival firms that are competing for each consumer. Besides the fast food industry, we believe our results also apply to other industries where chains hold a large share of the market, such as home improvement and department stores.

Finally, industries dominated by small retailers are also likely to satisfy all three conditions. The market for industries such as automotive dealers, plumbing contractors, household repairs, dentists, and veterinarians is composed by a large number of small firms. For example, in Illinois there are 7,300 licensed plumbers (McDevitt 2014). Consumers typically have little information regarding the firms that operate in the market. In these settings, word of mouth plays a major role in consumer awareness. Each consumer is likely to be aware of a small subset of the firms that operate in the market when making a purchase. Therefore,

[8] This is particularly true, for example, in the U.K. (Chintagunta et al. 2012), the Netherlands (Wildenbeest 2011), and Japan (Nishida 2015). Even in the U.S., where individual managers may have pricing discretion, local prices may be constrained by the prices consumers can obtain on the company's website (Hviid and Shaffer 2012). Even when prices of the stores in a chain vary from store to store, the longer-run decisions (such as the product line or the decision to offer a Price Matching Guarantee) are done at the chain level.

firms that operate in these markets are uncertain regarding the rivals that are competing with them for each consumer.

References

Anderson, S. P. and Renault, R. (1999). "Pricing, product diversity, and search costs: a Bertrand-Chamberlin-Diamond model". *Rand Journal of Economics*, 30:719–735.

Bachi, B. and Spiegler, R. (2018). Buridanic competition. *Games and Economic Behavior*, 107:298–315.

Biehal, G. and Chakravarti, D. (1986). Consumers' use of memory and external information in choice: Macro and micro perspectives. *Journal of Consumer Research*, 12(4):382–405.

Burdett, K. and Judd, K. L. (1983). Equilibrium price dispersion. *Econometrica: Journal of the Econometric Society*, pages 955–969.

Butters, G. R. (1977). Equilibrium distributions of sales and advertising prices. *The Review of Economic Studies*, 44(3):465–491.

Carlin, B. I. (2009). "Strategic price complexity in retail financial markets". *Journal of Financial Economics*, 91:278–287.

Champsaur, P. and Rochet, J.-C. (1989). Multiproduct duopolists. *Econometrica*, pages 533–557.

Chintagunta, P. K., Chu, J., and Cebollada, J. (2012). Quantifying transaction costs in online/off-line grocery channel choice. *Marketing Science*, 31(1):96–114.

Chioveanu, I. and Zhou, J. (2013). "Price competition with consumer confusion". *Management Science*, 59:2450–2469.

Deneckere, R. and Peck, J. (1995). Competition over price and service rate when demand is stochastic: A strategic analysis. *The RAND Journal of Economics*, pages 148–162.

Draganska, M. and Klapper, D. (2011). Choice set heterogeneity and the role of advertising: An analysis with micro and macro data. *Journal of Marketing Research*, 48(4):653–669.

Dubovik, A. and Janssen, M. C. (2012). Oligopolistic competition in price and quality. *Games and Economic Behavior*, 75(1):120–138.

Eliaz, K. and Spiegler, R. (2011). Consideration sets and competitive marketing. *The Review of Economic Studies*, 78(1):235–262.

Ellison, G. and Wolitzky, A. (2012). ”A search cost model of obfuscation”. *Rand Journal of Economics*, 43:417–441.

Fosfuri, A. and Giarratana, M. S. (2009). Masters of war: Rivals’ product innovation and new advertising in mature product markets. *Management Science*, 55(2):181–191.

Fudenberg, D. and Levine, D. K. (1988). Open-loop and closed-loop equilibria in dynamic games with many players. *Journal of Economic Theory*, 44(1):1–18.

Goeree, M. S. (2008). Limited information and advertising in the us personal computer industry. *Econometrica*, 76(5):1017–1074.

Gu, Y. and Wenzel, T. (2014). Strategic obfuscation and consumer protection policy. *Journal of Industrial Economics*, 62(4):632–660.

Hart, O. D. (1985). Monopolistic competition in the spirit of chamberlin: A general model. *The Review of Economic Studies*, 52(4):529–546.

Hauser, J. R. and Wernerfelt, B. (1990). An evaluation cost model of consideration sets. *Journal of consumer research*, 16(4):393–408.

Honka, E., Hortaçsu, A., and Vitorino, M. A. (2017). Advertising, consumer awareness, and choice: Evidence from the us banking industry. *The RAND Journal of Economics*, 48(3):611–646.

Houde, J.-F. (2012). Spatial differentiation and vertical mergers in retail markets for gasoline. *American Economic Review*, 102(5):2147–82.

Hviid, M. and Shaffer, G. (2012). Optimal low-price guarantees with anchoring. *Quantitative Marketing and Economics*, 10(4):393–417.

Ireland, N. J. (1993). The provision of information in a bertrand oligopoly. *The Journal of Industrial Economics*, pages 61–76.

Janssen, M. C. and Parakhonyak, A. (2013). Price matching guarantees and consumer search. *International Journal of Industrial Organization*, 31(1):1–11.

Laurent, G., Kapferer, J.-N., and Roussel, F. (1995). The underlying structure of brand awareness scores. *Marketing Science*, 14(3_supplement):G170–G179.

Lleras, J. S., Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2017). When more is less: Limited consideration. *Journal of Economic Theory*, 170:70–85.

Lynch Jr, J. G., Marmorstein, H., and Weigold, M. F. (1988). Choices from sets including remembered brands: Use of recalled attributes and prior overall evaluations. *Journal of Consumer Research*, 15(2):169–184.

Mamadehussene, S. (2019). Price-matching guarantees as a direct signal of low prices. *Journal of Marketing Research*, 56(2):245–258.

Mamadehussene, S. (2020a). The interplay between obfuscation and prominence in price comparison platforms. *Management Science*.

Mamadehussene, S. (2020b). "Measuring the competition effects of price matching guarantees". *mimeo*.

Manzini, P. and Mariotti, M. (2014). Stochastic choice and consideration sets. *Econometrica*, 82(3):1153–1176.

Masatlioglu, Y., Nakajima, D., and Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, 102(5):2183–2205.

McAfee, R. P. (1994). Endogenous availability, cartels, and merger in an equilibrium price dispersion. *Journal of Economic Theory*, 62(1):24–47.

McDevitt, R. C. (2014). "a" business by any other name: firm name choice as a signal of firm quality. *Journal of Political Economy*, 122(4):909–944.

Nedungadi, P. (1990). Recall and consumer consideration sets: Influencing choice without altering brand evaluations. *Journal of consumer research*, 17(3):263–276.

Nishida, M. (2015). Estimating a model of strategic network choice: The convenience-store industry in okinawa. *Marketing Science*, 34(1):20–38.

Perloff, J. M. and Salop, S. C. (1985). Equilibrium with product differentiation. *The Review of Economic Studies*, 52(1):107–120.

Piccione, M. and Spiegler, R. (2012). Price competition under limited comparability. *The quarterly journal of economics*, 127(1):97–135.

Robert, J. and Stahl, D. O. (1993). Informative price advertising in a sequential search model. *Econometrica: Journal of the Econometric Society*, pages 657–686.

Roberts, J. H. and Lattin, J. M. (1991). Development and testing of a model of consideration set composition. *Journal of Marketing Research*, 28(4):429–440.

Shaked, A. and Sutton, J. (1982). Relaxing price competition through product differentiation. *The review of economic studies*, pages 3–13.

Shugan, S. M. (1980). The cost of thinking. *Journal of consumer Research*, 7(2):99–111.

Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 69(1):99–118.

Stegeman, M. (1991). Advertising in competitive markets. *The American Economic Review*, pages 210–223.

Thisse, J.-F. and Vives, X. (1988). On the strategic choice of spatial price policy. *The American Economic Review*, pages 122–137.

Tsai, Y.-L. and Honka, E. (2018). Non-informational advertising informing consumers: How advertising affects consumers' decision-making in the us auto insurance industry. *Available at SSRN 3094448*.

Vandenbosch, M. B. and Weinberg, C. B. (1995). Product and price competition in a two-dimensional vertical differentiation model. *Marketing Science*, 14(2):224–249.

Varian, H. R. (1980). A model of sales. *The American Economic Review*, 70(4):651–659.

Wiggins, S. N. and Lane, W. J. (1983). Quality uncertainty, search, and advertising. *The American Economic Review*, 73(5):881–894.

Wildenbeest, M. R. (2011). An empirical model of search with vertically differentiated products. *The RAND Journal of Economics*, 42(4):729–757.

Wilson, C. M. (2010). "Ordered search and equilibrium obfuscation". *International Journal of Industrial Organization*, 28:496–506.

Appendix

A Solution to the sequential game from the example in Section 2.2

Throughout this appendix we denote by $\pi_0(r)$ ($\pi_1(r)$) the profit of a firm that played $\alpha = 0$ ($\alpha = 1$) in the subgame where exactly r firms play $\alpha = 0$. We start by solving the second stage of the game. In the analysis of each subgame we denote by F_0 (F_1) the price distribution of a firm that played $\alpha = 0$ ($\alpha = 1$).

Subgame r=0

This subgame is similar to Varian(1980). All firms earn profit $\frac{\lambda}{2}$, i.e. $\pi_1(r = 0) = \frac{\lambda}{2}$

Subgame r=1

By a standard argument, both price distributions must have the same lower bound, which we denote by $\underline{p}$. Moreover, the upper bound of F_1 is the consumer reservation value, and F_1 can have no mass points. It follows that $\pi_1(r = 1) = \frac{1}{N-1}\lambda + \frac{N-2}{N-1}\frac{\lambda}{2} = \frac{\lambda}{2}\frac{N}{N-1}$.

When a firm with $\alpha = 1$ plays $\underline{p}$ it obtains profit $\left(\frac{\lambda}{2}\frac{N}{N-1} + 1 - \lambda\right)$. Therefore, it follows that $\underline{p} = \frac{\frac{\lambda}{2}\frac{N}{N-1}}{\frac{\lambda}{2}\frac{N}{N-1}+1-\lambda}$. This then implies that $\pi_0(r = 1) = (1 - \lambda)\underline{p} = \frac{(1-\lambda)\frac{\lambda}{2}\frac{N}{N-1}}{\frac{\lambda}{2}\frac{N}{N-1}+1-\lambda}$

Subgame $2 \leq r \leq N - 2$

By an argument similar to Varian(1980), it follows that neither F_0 nor F_1 have mass points. Morevoer, the upper bound of F_1 is at the consumer reservation value. This implies that

$$\pi_1(r) = \frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right] \tag{5}$$

Let $\underline{p}_0$ and $\underline{p}_1$ denote the lower bounds of F_0 and F_1.

Lemma 2 $\underline{p}_0 \leq \underline{p}_1$.

Proof. Suppose not, i.e. $\underline{p}_1 < \underline{p}_0$. This implies that

$$\pi_0(\underline{p}_1) = (1-\lambda)\underline{p}_1$$

$$\pi_0(\underline{p}_0) = (1-\lambda)\left[\frac{r-1}{N-1}\underline{p}_0 + \frac{N-r}{N-1}[1-F_1(\underline{p}_0)]\underline{p}_0\right]$$

$$\pi_1(\underline{p}_0) = \frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]\underline{p}_0 + (1-\lambda)\left[\frac{r}{N-1}\underline{p}_0 + \frac{N-r-1}{N-1}[1-F_0(\underline{p}_0)]\underline{p}_0\right]$$

$$\pi_1(\underline{p}_1) = \frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]\underline{p}_1 + (1-\lambda)\underline{p}_1$$

It follows that

$$\begin{aligned}
\pi_0(\underline{p}_1) - \pi_0(\underline{p}_0) &= (1-\lambda)\left[\underline{p}_1 - \frac{r-1}{N-1}\underline{p}_0 + \frac{N-r}{N-1}[1-F_1(\underline{p}_0)]\underline{p}_0\right] \\
&> (1-\lambda)\left[\underline{p}_1 - \frac{r}{N-1}\underline{p}_0 + \frac{N-r-1}{N-1}[1-F_1(\underline{p}_0)]\underline{p}_0\right] \\
&> (1-\lambda)\left[\underline{p}_1 - \frac{r}{N-1}\underline{p}_0 + \frac{N-r-1}{N-1}[1-F_1(\underline{p}_0)]\underline{p}_0\right] - \frac{\lambda}{2}\left[1+\frac{r}{N-1}\right](\underline{p}_0 - \underline{p}_1) \\
&= \pi_1(\underline{p}_1) - \pi_1(\underline{p}_0) \\
&= 0
\end{aligned}$$

which contradicts that $\underline{p}_0$ is in the support of F_0. □

Lemma 3 *If* $r \in [2, N-2]$ *is an equilibrium of the first stage, then* $\frac{1-\lambda}{N-1} > \frac{\lambda}{2}\left[1+\frac{r+1}{N-1}\right]$

Proof. From before, we have that $\pi_1(r) = \frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]$. When a firm with $\alpha = 1$ plays $\underline{p}_0$ it makes profit $\left(\frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]+1-\lambda\right)\underline{p}_0$. It then follows that $\underline{p}_0 \leq \frac{\frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]}{\frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]+1-\lambda}$. Therefore

$$\pi_0 = (1-\lambda)\underline{p}_0 \leq \frac{(1-\lambda)\frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]}{\frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]+1-\lambda}$$

In equilibrium of the first stage, it must be that $\pi_0(r) \geq \pi_1(r-1)$. After some algebra this condition becomes $\frac{2(1-\lambda)}{\lambda} \geq N+r+r-2+\frac{r(r-1)}{N-1}$. Because $r \geq 2$, this implies that $\frac{2(1-\lambda)}{\lambda} > N+r$, which is equivalent to $\frac{1-\lambda}{N-1} > \frac{\lambda}{2}\left[1+\frac{r+1}{N-1}\right]$. □

Lemma 4 *If* $2 \leq r \leq N-2$ *and* $\frac{1-\lambda}{N-1} > \frac{\lambda}{2}\left[1+\frac{r+1}{N-1}\right]$, *then* $\underline{p}_0 = \underline{p}_1$.

Proof. Suppose, by contradiction, that $\underline{p}_0 \neq \underline{p}_1$. Then it follows from Lemma 2 that $\underline{p}_0 < \underline{p}_1$. Let ϵ be small enough such that $\underline{p}_1 - \epsilon > \underline{p}_0$.

$$\pi_0(\underline{p}_1) = (1-\lambda)\left[\frac{r-1}{N-1}[1-F_0(\underline{p}_1)]\underline{p}_1 + \frac{N-r}{N-1}\underline{p}_1\right]$$

$$\pi_0(\underline{p}_1-\epsilon) = (1-\lambda)\left[\frac{r-1}{N-1}[1-F_0(\underline{p}_1-\epsilon)](\underline{p}_1-\epsilon) + \frac{N-r}{N-1}(\underline{p}_1-\epsilon)\right]$$

$$\pi_1(\underline{p}_1) = \frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]\underline{p}_1 + (1-\lambda)\left[\frac{r}{N-1}[1-F_0(\underline{p}_1)]\underline{p}_1 + \frac{N-1-r}{N-1}\underline{p}_1\right]$$

$$\pi_1(\underline{p}_1-\epsilon) = \frac{\lambda}{2}\left[1+\frac{r}{N-1}\right](\underline{p}_1-\epsilon)+(1-\lambda)\left[\frac{r}{N-1}[1-F_0(\underline{p}_1-\epsilon)](\underline{p}_1-\epsilon)+\frac{N-1-r}{N-1}(\underline{p}_1-\epsilon)\right]$$

Let $\Delta \equiv [1-F_0(\underline{p}_1-\epsilon)](\underline{p}_1-\epsilon) - [1-F_0(\underline{p}_1)]\underline{p}_1$.

$$\pi_0(\underline{p}_1-\epsilon) - \pi_0(\underline{p}_1) = (1-\lambda)\left[\frac{r-1}{N-1}\Delta - \frac{N-r}{N-1}\epsilon\right]$$

Notice that because both $\underline{p}_1$ and $\underline{p}_1-\epsilon$ are in the support of F_0, it follows that $\pi_0(\underline{p}_1-\epsilon) =$

$\pi_0(\underline{p}_1)$ and, therefore, $\Delta > 0$.

$$\begin{aligned}
\pi_1(\underline{p}_1 - \epsilon) - \pi_1(\underline{p}_1) &= -\frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right]\epsilon + (1-\lambda)\left[\frac{r}{N-1}\Delta - \frac{N-r-1}{N-1}\epsilon\right] \\
&= -\frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right]\epsilon + (1-\lambda)\frac{\Delta+\epsilon}{N-1} + \pi_0(\underline{p}_1 - \epsilon) - \pi_0(\underline{p}_1) \\
&= -\frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right]\epsilon + (1-\lambda)\frac{\Delta+\epsilon}{N-1} \\
&> \left(\frac{1-\lambda}{N-1} - \frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right]\right)\epsilon \\
&> \left(\frac{1-\lambda}{N-1} - \frac{\lambda}{2}\left[1 + \frac{r+1}{N-1}\right]\right)\epsilon \\
&> 0
\end{aligned}$$

which contradicts that $\underline{p}_1$ is in the support of F_1. □

Therefore, we find that if $2 \leq r \leq N-2$ and $\frac{1-\lambda}{N-1} > \frac{\lambda}{2}\left[1 + \frac{r+1}{N-1}\right]$, then F_0 and F_1 have the same lower bound. Let us denote such lower bound by $\underline{p}$. The profit of a firm with $\alpha = 1$ when it plays $\underline{p}$ is $\left(\frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right] + 1 - \lambda\right)\underline{p}$. Using (5), it follows that $\underline{p} = \frac{\frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]}{\frac{\lambda}{2}\left[1+\frac{r}{N-1}\right]+1-\lambda}$. It then follows that $\pi_0(r) = (1-\lambda)\underline{p}$ or, equivalently:

$$\pi_0(r) = \frac{(1-\lambda)\frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right]}{\frac{\lambda}{2}\left[1 + \frac{r}{N-1}\right] + 1 - \lambda} \tag{6}$$

Subgame r=N-1

Lemma 5 $\pi_1(r = N-1) = \lambda$. *If* $N > 2 + \frac{1-\lambda}{\lambda}$ *then* $\pi_0(r = N-1) = \frac{1-\lambda}{N-1}$, *otherwise* $\pi_0(r = N-1) = \lambda(1-\lambda)$.

Proof. By a standard argument, the upper bound of both F_0 and F_1 is the consumer reservation value. Moreover, F_0 has no mass points. This implies that $\pi_1 = \lambda$.

It can be shown that if $N > 2 + \frac{1-\lambda}{\lambda}$ then in equilibrium F_1 is degenerate at $p = 1$.

Therefore, $\pi_0 = \pi_0(p = 1) = \frac{1-\lambda}{N-1}$. If, instead, $N \leq 2 + \frac{1-\lambda}{\lambda}$ it can be show that there is an equilibrium under which $\underline{p}_0 = \underline{p}_1$. In this case, because $\pi_1(\underline{p}_1) = \underline{p}_1$, it follows that $\underline{p}_1 = \lambda$ and, therefore, $\pi_0 = \lambda(1 - \lambda)$. □

Subgame r=N

This is a Bertrand game, in which all firms make zero profit.

Solving the first-stage

Proof of Proposition 1. $r^* = 0$ if $\pi_1(0) \geq \pi_0(1)$. After some algebra the condition simplifies to $N \geq \frac{2(1-\lambda)}{\lambda}$. □

It is straightforward to see that $r = N$ is not an equilibrium.

In order for some $1 \leq r \leq N - 1$ to be an equilibrium, the following two conditions must hold:

i) $\pi_0(r) \geq \pi_1(r - 1)$

ii) $\pi_1(r) \geq \pi_0(r + 1)$

After some algebra, these conditions simplify to $t \leq r \leq 1 + t$, where

$$t = \frac{1}{2}\left(1 - 2N + \sqrt{\frac{8(N-1)(1-\lambda)}{\lambda} + 1}\right) \tag{7}$$

Proof of Proposition 2

When all firms play $\alpha = 1$, they play price distribution F_{Sim} in the second stage, and each firm earns profit $\frac{\lambda}{2}$. We will show that no firm benefits from deviating to $\alpha = 0$ in the first stage. Because firms are symmetric, it suffices to examine the subgame in which one firm (say N_1) plays $\alpha = 0$ whereas the remaining firms play $\alpha = 1$.

In the second stage, firms S_1 and S_2 are symmetric. We denote their price distribution by F_S. We denote by F_L and F_H the price distributions of N_1 and N_2. Finally, we denote the respective lower bounds of these price distributions by $\underline{p}_S$, $\underline{p}_L$ and $\underline{p}_H$.

Lemma 6 $\underline{p}_S = \underline{p}_L$.

Proof. We prove this by contradiction. We consider two cases:

Case 1: $\underline{p}_S < \underline{p}_L$

$$\pi_L(\underline{p}_L) = (1-\lambda)\left[\left(\frac{1}{3}+\epsilon\right)[1-F_H(\underline{p}_L)]\underline{p}_L + \left(\frac{2}{3}-\epsilon\right)[1-F_S(\underline{p}_L)]\underline{p}_L\right]$$

$$\pi_L(\underline{p}_S) = (1-\lambda)\underline{p}_S$$

$$\pi_S(\underline{p}_L) = \frac{\lambda}{2}\left[\frac{4}{3}-\frac{\epsilon}{2}\right]\underline{p}_L+(1-\lambda)\left[\left(\frac{1}{3}-\frac{\epsilon}{2}\right)[1-F_H(\underline{p}_L)]\underline{p}_L+\left(\frac{1}{3}-\frac{\epsilon}{2}\right)\underline{p}_L+\left(\frac{1}{3}+\epsilon\right)[1-F_S(\underline{p}_L)]\underline{p}_L\right]$$

$$\pi_S(\underline{p}_S) = \frac{\lambda}{2}\left[\frac{4}{3}-\frac{\epsilon}{2}\right]\underline{p}_S + (1-\lambda)\underline{p}_S$$

After some algebra it follows that $\pi_L(\underline{p}_S) - \pi_L(\underline{p}_L) > \pi_S(\underline{p}_S) - \pi_S(\underline{p}_L)$, which contradicts that $\underline{p}_L$ is in the support of F_L.

Case 2: $\underline{p}_L < \underline{p}_S$

Using a standard argument, in this case it must be that $\underline{p}_L = \underline{p}_H$.

$$\pi_S(\underline{p}_H) = \frac{\lambda}{2}\left[\frac{4}{3}-\frac{\epsilon}{2}\right]\underline{p}_H + (1-\lambda)\underline{p}_H$$

$$\pi_S(\underline{p}_S) = \frac{\lambda}{2}\left[\frac{4}{3}-\frac{\epsilon}{2}\right]\underline{p}_S+(1-\lambda)\left[\left(\frac{1}{3}+\epsilon\right)\underline{p}_S+\left(\frac{1}{3}-\frac{\epsilon}{2}\right)[1-F_H(\underline{p}_S)]\underline{p}_S+\left(\frac{1}{3}-\frac{\epsilon}{2}\right)[1-F_L(\underline{p}_S)]\underline{p}_S\right]$$

$$\pi_H(\underline{p}_H) = \frac{\lambda}{2}\left[\frac{4}{3}+\epsilon\right]\underline{p}_H + (1-\lambda)\underline{p}_H$$

$$\pi_H(\underline{p}_S) = \frac{\lambda}{2}\left[\frac{4}{3}+\epsilon\right]\underline{p}_S + (1-\lambda)\left[\left(\frac{2}{3}-\epsilon\right)\underline{p}_S + \left(\frac{1}{3}+\epsilon\right)[1-F_L(\underline{p}_S)]\underline{p}_S\right]$$

After some algebra it follows that $\pi_S(\underline{p}_H)-\pi_S(\underline{p}_S) > \pi_H(\underline{p}_H)-\pi_H(\underline{p}_S)$, which contradicts that $\underline{p}_S$ is in the support of F_S. □

Let $\underline{p} \equiv \underline{p}_S = \underline{p}_L$.

Lemma 7 $\underline{p}_H = \underline{p}$

Proof. By a standard argument, it must be that $\underline{p}_H \geq \underline{p}$. Suppose, by contradiction, that $\underline{p}_H \neq \underline{p}$. This would imply that $\underline{p}_H > \underline{p}$.

$$\pi_H(\underline{p}) = \frac{\lambda}{2}\left[\frac{4}{3}+\epsilon\right]\underline{p} + (1-\lambda)\underline{p}$$

$$\pi_H(\underline{p}_H) = \frac{\lambda}{2}\left[\frac{4}{3}+\epsilon\right]\underline{p}_H + (1-\lambda)\left[\left(\frac{1}{3}+\epsilon\right)[1-F_L(\underline{p}_H)]\underline{p}_H + \left(\frac{2}{3}-\epsilon\right)[1-F_S(\underline{p}_H)]\underline{p}_H\right]$$

$$\pi_S(\underline{p}) = \frac{\lambda}{2}\left[\frac{4}{3}-\frac{\epsilon}{2}\right]\underline{p} + (1-\lambda)\underline{p}$$

$$\pi_S(\underline{p}_H) = \frac{\lambda}{2}\left[\frac{4}{3}-\frac{\epsilon}{2}\right]\underline{p}_H+(1-\lambda)\left[\left(\frac{1}{3}-\frac{\epsilon}{2}\right)[1-F_L(\underline{p}_H)]\underline{p}_H+\left(\frac{1}{3}+\epsilon\right)[1-F_S(\underline{p}_H)]\underline{p}_H+\left(\frac{1}{3}-\frac{\epsilon}{2}\right)\underline{p}_H\right]$$

$$\pi_L(\underline{p}) = (1-\lambda)\underline{p}$$

$$\pi_L(\underline{p}_H) = (1-\lambda)\left[\left(\frac{2}{3}-\epsilon\right)[1-F_S(\underline{p}_H)]\underline{p}_H + \left(\frac{1}{3}+\epsilon\right)\underline{p}_H\right]$$

Because $\underline{p}$ and $\underline{p}_H$ are both on the support of F_L and F_S, it follows that:

$$\pi_L(\underline{p}) = \pi_L(\underline{p}_H) \iff \underline{p}_H - \underline{p} = \left(\frac{2}{3} - \epsilon\right) F_S(\underline{p}_H)\underline{p}_H \tag{8}$$

$$\pi_S(\underline{p}) = \pi_S(\underline{p}_H) \iff \left(\frac{1}{3} - \frac{\epsilon}{2}\right) F_L(\underline{p}_H)\underline{p}_H = -M F_S(\underline{p}_H)\underline{p}_H \tag{9}$$

where $M = \left(\frac{2}{3} - \epsilon\right)\left(\frac{\lambda}{2(1-\lambda)}\left[\frac{4}{3} - \frac{\epsilon}{2}\right] + 1\right) + \frac{1}{3} + \epsilon$

After some algebra it follows that $\pi_H(\underline{p}) - \pi_H(\underline{p}_H) > \pi_S(\underline{p}) - \pi_S(\underline{p}_H)$ provided that the following condition holds:

$$\frac{1}{3} > \epsilon\left[\frac{3}{4}\frac{\lambda}{1-\lambda}\left(\frac{2}{3} - \epsilon\right) + 2 + \frac{3}{2\left(\frac{1}{3} - \frac{\epsilon}{2}\right)} M\right]$$

The above condition holds for $\epsilon < 0.04$, which contradicts that $\underline{p}_H$ is in the support of F_H. $\square$

The upper bound of F_H is the consumer reservation value, and neither F_S nor F_L can have a mass point at such price. It then follows that $\pi_H = \frac{\lambda}{2}\left[\frac{4}{3} + \epsilon\right]$. In order for $\underline{p}$ to be on the support of F_H, it must be that

$$\underline{p} = \frac{\frac{\lambda}{2}\left[\frac{4}{3} + \epsilon\right]}{\frac{\lambda}{2}\left[\frac{4}{3} + \epsilon\right] + 1 - \lambda}$$

The profit of firm N_1 when it plays price $\underline{p}$ is $(1 - \lambda)\underline{p}$. It follows that

$$\pi_L = \frac{(1-\lambda)\frac{\lambda}{2}\left[\frac{4}{3} + \epsilon\right]}{\frac{\lambda}{2}\left[\frac{4}{3} + \epsilon\right] + 1 - \lambda}$$

In the first stage, firm N_1 will prefer to play $\alpha = 1$ if $\frac{\lambda}{2} \geq \frac{(1-\lambda)\frac{\lambda}{2}\left[\frac{4}{3}+\epsilon\right]}{\frac{\lambda}{2}\left[\frac{4}{3}+\epsilon\right]+1-\lambda} \iff \epsilon < \frac{1}{6}\frac{3\lambda-1}{2-3\lambda}$.

Because $\frac{3\lambda-1}{2-3\lambda}$ is increasing in λ, a sufficient condition is that the above inequality holds at $\lambda = 0.4$. Thus, all firms playing $\alpha = 1$ in the first stage is an equilibrium provided that $\epsilon < \frac{1}{6}\frac{3\cdot 0.4-1}{2-3\cdot 0.4} \approx 0.0417$.

B Proofs for section 3

B.1 Proof of proposition 3

Due to symmetry we omit the subscript i on functions b^i and c^i. A symmetric strategy profile is such that each player follows the same (mixed) strategy. Let $\sigma(n) \in \Delta Z$ be such $(\sigma(n), ..., \sigma(n)) \equiv s(n)$ is a symmetric Nash equilibrium of the simultaneous game when $N = n$. Because firms are symmetric and there is a consumer per firm, it follows that

$$b(z_i, z_{-i}) = \frac{N}{n}\frac{1}{C_n^N}\int_{\gamma\in\Gamma} \mathbf{1}\{i \in \gamma\} v(z_i, z_{\gamma-i}) d\gamma.$$

Suppose there is some $N > n \geq 2$ such that $s(n)$ is not a Nash equilibrium of the simultaneous game. Then, there must be a firm i and some $z_i' \in Z$ such that

$$E_{z_{-i}}(b(z_i', z_{-i}) | s_{-i}(n)) - c(z_i') > E_{z_i,z_{-i}}(b(z_i, z_{-i}) | s(n)) - E(c(z_i') | s(n)) \quad (10)$$

Notice that for any $z_i \in Z$,

$$E_{z_{-i}}(b(z_i, z_{-i}) | s_{-i}(n)) = \frac{N}{n}\frac{n}{N}E_{z_{-i}}(v(z_i, z_{-i}) | s_{-i}(n))$$

Therefore, (10) is equivalent to

$$E_{z_{-i}}\left(v\left(z_i', z_{-i}\right)|s_{-i}\left(n\right)\right) - c\left(z_i'\right) > E_{z_i, z_{-i}}\left(b\left(z_i, z_{-i}\right)|s\left(n\right)\right) - E\left(c\left(z_i'\right)|s\left(n\right)\right),$$

which contradicts $s\left(n\right)$ being a symmetric Nash equilibrium of the simultaneous game when $N = n$.

B.2 Proof of proposition 4

The argument essentially follows \cite{fudenberg1988open}. Take any Nash equilibrium $\sigma : T \to \Delta Z$ of the simultaneous game. For all $t \in T$, let $\sigma_x^t \in \Delta X$ be such that

$$\sigma_x^t\left(x_i\right) = \int\limits_{y_i \in Y} \sigma^t\left(x_i, y_i\right) dy_i$$

for all $x_i \in X_i$. For all $f_x : T \to \left(\Delta X\right)$ such that $f_x =^{a.e.} \sigma_x$, let $\sigma_y^t\left(x_i, f_x\right)$ be such that

$$\sigma_y^t\left(x_i, f_x\right)\left(y_i\right) = \frac{\sigma^t\left(x_i, y_i\right)}{\int\limits_{\widehat{y}_i \in Y} \sigma^t\left(x_i, \widehat{y}_i\right) d\widehat{y}_i}$$

for all $y_i \in Y_i$, for all $x_i \in supp\left(f_x^t\right)$ and for all $t \in T$. For all other $f_x : T \to \left(\Delta X\right)$, let $\sigma_y\left(\cdot, f_x\right)$ be anything that satisfies condition ii), which is assumed to exist. We show that $\left(\sigma_x, \sigma_y\right)$ is a SPE of the sequential game, which, by design, implies that σ is a SPE outcome of the sequential game.

First, notice that condition i) of the definition is satisfied whenever $f_x =^{a.e.} \sigma_x$, because, if not, there would be some $t \in T$, some $x_i' \in supp\left(f_x^t\right)$ and some $y_i' \in Y_i$ such that

$$\pi^t\left(\left(x_i', y_i'\right), \left(\sigma_x, g_y\left(\sigma_x, \sigma_y\right)\right)\right) > \int\limits_{y_i \in Y} \sigma_y^t\left(x_i, \sigma_x\right)\left(y_i\right) \pi^t\left(\left(x_i', y_i\right), \left(\sigma_x, g_y\left(\sigma_x, \sigma_y\right)\right)\right) dy_i$$

which implies that

$$\pi^t\left(z_i', \sigma\right) > \int\limits_{z_i \in Z} \sigma^t\left(z_i\right) \pi^t\left(z_i, \sigma\right) dz_i$$

where $z_i' = (x_i', y_i')$, which is a contradiction. Condition i) is also satisfied for all other $f_x : T \to (\Delta X)$ by design.

Condition ii) must also be satisfied, because, if not, there would be some $t \in T$ and some (x_i', y_i') such that

$$\begin{aligned}
&\pi^t\left(\left(x_i', y_i'\right), \left(\sigma_x, g_y\left(\sigma_x, \sigma_y\right)\right)\right) \\
&\geq \int\limits_{y_i \in Y} \sigma_y^t\left(x_i', \sigma_x\right)\left(y_i\right) \pi^t\left(\left(x_i', y_i\right), \left(\sigma_x, g_y\left(\sigma_x, \sigma_y\right)\right)\right) dy_i \\
&> \int\limits_{x_i \in X} \int\limits_{y_i \in Y} \sigma_x^t\left(x_i\right) \sigma_y^t\left(x_i, \sigma_x\right)\left(y_i\right) \pi^t\left(\left(x_i, y_i\right), \left(\sigma_x, g_y\left(\sigma_x, \sigma_y\right)\right)\right) dy_i dx_i \\
&= \int\limits_{z_i \in Z} \sigma^t\left(z_i\right) \pi^t\left(z_i, \sigma\right) dz_i
\end{aligned}$$

which is a contradiction. Seeing as conditions i) and ii) are satisfied, profile (σ_x, σ_y) is a SPE of the sequential game.

Now consider a SPE (σ_x, σ_y) and the corresponding SPE outcome $\sigma : T \to \Delta Z$ of the sequential game. We show that it is also a Nash equilibrium of the simultaneous game. Take any $t \in T$ and any $z_i' = (x_i', y_i') \in Z$. It follows that

$$\begin{aligned}
&\int\limits_{z_i \in Z} \sigma^t(z_i)\, \pi^t(z_i, \sigma)\, dz_i \\
&= \int\limits_{x_i \in X} \int\limits_{y_i \in Y} \sigma_x^t(x_i)\, \sigma_y^t(x_i, \sigma_x)(y_i)\, \pi^t\left((x_i, y_i), (\sigma_x, g_y(\sigma_x, \sigma_y))\right) dy_i dx_i \\
&\geq \int\limits_{y_i \in Y} \sigma_y^t(x_i', \sigma_x)(y_i)\, \pi^t\left((x_i', y_i), (\sigma_x, g_y(\sigma_x, \sigma_y))\right) dy_i \\
&\geq \pi^t\left((x_i', y_i'), \sigma\right).
\end{aligned}$$

where the first inequality follows by condition ii), while the second inequality follows by condition i), which proves that σ is a Nash equilibrium of the simultaneous game.