

Stackelberg versus Cournot Oligopoly with Private Information

Eray Cumbul^{*†}

IESE Business School, Barcelona

Abstract

In this paper, we compare an n -firm Cournot game with a Stackelberg model, where n -firms choose outputs sequentially, in a stochastic demand environment with private information. The Stackelberg perfect revealing equilibrium expected output and total surplus are lower while expected price and total profits are higher than the Cournot equilibrium ones irrespective of how noisy both the demand shocks and private demand signals of firms are. These rankings are the opposite to the rankings of prices, total output, surplus, and profits under perfect information. Our Stackelberg model identifies the presence of four effects, which are absent under the Cournot model. Because of *i) the signaling effect*, early-mover firms would like to set low quantities to signal to their followers that the demand is low. This effect reduces *ii) first-mover advantages*. Moreover, as followers infer the demand signals of their predecessors, they are better informed about demand compared to Cournot oligopolists. But this *iii) information acquisition* of followers also imposes *iv) negative externalities* on their rivals as rivals have less value from exploiting their demand information. Only *i)* and *iv)* favor Cournot over Stackelberg in welfare terms and they are the dominant ones. We also study a number of implications of our results in examining the relationships of prices, profits, and welfare with market concentration.

^{*}ecumbul@iese.edu, This version: March 30, 2014

[†]I am deeply grateful to Gábor Virág and Xavier Vives for their guidance and patience. I would like to thank seminar participants at the IESE Business School, Bilkent University, TOBB University of Economics and Technology, ADA university, Antalya International University, University of Rochester, University of Toronto St. George, and University of Toronto Mississauga. I also thank Paulo Barelli, Andrew F. Daughety, Alberto Galasso, Srihari Govindan, Asen Kochov, Semih Koray, Martin Osborne, Morten Olsen, Ayca Ozdogan, Roman Pancs, Carolyn Pitchik, Michael Raith, William Thomson, Mihkel M. Tombak, and Kemal Yildiz for their useful comments and suggestions. I thank Tubitak for their financial support. Of course, all errors are mine.

1 Introduction

The Stackelberg- and the Cournot models are two of the leading frameworks in which economists have studied oligopolistic competition. Among others, Anderson and Engers (1992) have argued that the simultaneous-move Cournot model is applicable to characterize an industry where lags in the observation of output decisions are long, whereas the sequential-move Stackelberg model applies when the reverse holds. While many industries fit the Cournot framework better¹, Shinkai (2000) has argued that the DRAM market (i.e., the market for the main memory component of most computers and many electronic systems) is better described by the Stackelberg model because firms make sequential capacity choices. Typically, manufacturers produce highly substitutable products and they need to make huge investments in an irreversible manner at least in the short-run in this market.^{2, 3}

As a second example for a Stackelberg industry, consider that broadband technology is about to enter a telecommunication market (De Miguel and Xu, 2009). The new entrants (i.e., leaders) must decide whether to invest in expanding their network capacity (or install a network) and offer the new service to the market. Since the network expansion process is time consuming, these companies have to make a decision in advance and thus realize the demand for the new technology only in distribution. However, the incumbent firms (i.e., followers) have the flexibility to wait until they observe the leaders' supply decisions to infer signals about the potential demand before deciding how much of their network capacity to allocate to the new service. Following the same example, the Cournot competition is likely to be associated with the situation where the new entrants are allowed to connect their equipment to the incumbents' network directly. This type of regulation is referred to as "*interconnection (collocation)*." The government may intervene the market by adopting this type of public promotions policies in the diffusion of broadband networks by providing subsidies (Gentzoglani and Ara-

¹For instance, Lopex et al. (2002) show that the Cournot competition is widespread in 32 US food processing industries over the 1972-1992 period.

²Another example of an industry in which the credible-quantity commitment assumption is valid would be the alumina-refining industry. (Ghemawat and Nalebuff, 1985)

³Kadiyali et al. (2001) point out three structural new empirical industrial organization (NEIO) approaches that can identify competitive interactions among firms (e.g., Cournot or Stackelberg) in a given industry. These approaches are 1) The menu approach, 2) The conjectural variations (CV) approach, and 3) The conduct parameter and the weighted profit approaches.

vantinos, 2008). As a policy implication, if the Cournot competition creates more total output than the Stackelberg competition (as in our paper), then this implies that broadband deployment is higher under the interconnection regulations than under non-interconnection. Such a relation would be consistent with many empirical findings in OECD countries and European Union⁴. Similarly, it has been argued that “*unbundling*” regulations, which allow entrants to have access to the incumbent’s facilities, promote broadband adoption in France, Japan, and South Korea (Wallsten, 2007).

We argue that the first-mover firm in a Stackelberg game tends to be the biggest player in industries where there is a fairly stable and predictable demand. For example, in the copying-machine industry, Xerox has always been the top since it introduced the world’s first plain-paper copier in 1959 (Liu, 2005). But in some markets whose demand is rather unpredictable, such as the markets of DRAM, first generation PC’s and medical devices and drugs, it seems that there are late-mover advantages. This is because the followers can wait and see customer’s response to a new product introduced by the first movers, as well as move along the learning curve of innovation (Liu, 2005). For instance, during the 1970s, Intel led the DRAM market (Kang, 2010). After the 1980s, new emerging Japanese conglomerates such as Hitachi and Mitsubishi took a chance to lead the market. Following its entry, the South-Korean firm Samsung took the lead in 1992 when competing with three major competitors Hynix, Micron and Infineon. It has continued to lead the market since then.⁵ A second example of a successful follower is in the medical device market. Boston-scientific (a global manufacturer and supplier of medical devices) claimed that it seizes 70% market share of drug-coated stents just one year after Johnson & Johnson first introduced this product (Abelson, 2004). These real-life observations are consistent with Cumbul (2014)’s recent theoretical finding that in a market with n firms ($n \leq 4$), the last mover firm always gets the highest profit in the same Stackelberg model studied in this paper. However, if the number of firms is greater than four, then the leader can still get the highest profit especially when firms have sufficiently low demand uncertainties. This can also partially explain Samsung’s first-mover advantages in

⁴For reference, we refer to Flamm (2005), Distaso et al. (2006) and Wallsten (2007).

⁵As of May 2012, the top five players that have approximately 91% of the global mobile DRAM market shares are Samsung (41.4 %), Hynix (23.9%), Elpida (12.4%), Micron (11.6%), and Winbond (1.7%) (DrameXchange, 2012).

the DRAM market based on its technological leadership starting from 1992 until today as not only the number of followers of Samsung increase⁶ but also demand uncertainties might have relatively decreased compared to the early stages of the product development.

Given the real life examples above, it is apparent that the follower firms might enjoy late-mover advantages by learning about demand from first-mover firms' strategies in industries where firms face demand uncertainties. In this paper, we want to compare Cournot and Stackelberg models with respect to total output, welfare and producer surplus in such an environment. This comparison is important for at least three reasons. First, such an understanding provides insights into the mechanics of those important theoretical models. Relatedly, it also helps us in deciding which framework (if either) is more appropriate for studying a given industry given the observed price and output levels. Second, once it has been decided which model fits a given industry better, the policy maker can assess better whether mergers or other industry developments may help or hurt consumers. The answer may very well depend on which model one thinks is more appropriate to describe an industry. Third, comparing these models also have implications on the relationship between market concentration and welfare, which is affected by the asymmetries among firms. When all firms are identical except for the timing of production (as in our model), the Stackelberg model yields a higher HHI (Herfindahl-Hirschman Index), which indicates a higher degree of concentration of the market, than does the Cournot model.⁷ Hence total welfare comparison between these models helps us understand whether concentration is beneficial or not for the society.

Unfortunately, the current literature compares the Stackelberg and Cournot equilibrium outcomes under the assumption that the demand is known.⁸ An important lesson learned is that total output and welfare under the Stackelberg competition are greater than or equal to the ones under the Cournot competition when the demand is known by the firms and costs are symmetric and linear. In

⁶Many Taiwanese firms such as Winbond, Nan ya, Promos, and Powerchip entered into the market following Samsung's success. (Kang, 2010)

⁷Besides firm asymmetries, there are other factors that can affect the degree of market concentration, e.g., the number of firms in the market. However, those factors are not in the scope of this paper.

⁸Please visit the related literature section for more discussion about these papers.

a leader-follower game with perfect information, the leader typically produces a larger quantity and makes larger profits whereas the follower produces a lower quantity and makes lower profits than in a simultaneous-move game. Strategic substitutability among firm quantity decisions is the main driving force behind this result. Therefore, as we move from simultaneous-move games to sequential-move ones, this reallocation of firms' output decisions increases total output, consumer surplus and total welfare, while it decreases total profits.

We argue that the above output, profit, and welfare rankings between Cournot and Stackelberg competitions, which can be found in most microeconomics textbooks, are reversed in a world of incomplete information about demand. To obtain these results, we compare an n -firm Cournot game with a Stackelberg model, where n firms choose outputs sequentially, in a stochastic demand environment with private information. Demand is linear and stochastic in the intercept. Firms have private information about the state of the demand. We assume prior and posterior distributions that generate posterior expected values that are linear in the observable signals. In that regard, our model will be quite general because it accommodates a rich class of distributions.⁹ In this model, we show that the Stackelberg perfect revealing equilibrium expected price is higher, so expected output is lower than in Cournot equilibrium ones for any finite number of firms. Although the same ranking also holds in terms of expected consumer surplus and total surplus comparisons, Stackelberg performs better than Cournot in expected total profit comparisons up to five firms.¹⁰ Moreover, our results hold regardless of how (symmetrically) noisy the private demand signals of firms are. In particular, they hold when the noise converges to zero. Therefore, our results also imply that the first-mover advantage of the leader is reduced if there is even a slight noise in the observation of demand.¹¹

In our proposed Stackelberg perfect revealing equilibrium, firms' optimal decisions might involve both strategic substitutability and complementarity relationships. We discuss that while the quantity strategy of any follower firm is a

⁹For example, the prior-posterior distribution pairs that satisfy this linearity assumption would include the Gamma-Poisson, Beta-Binomial, and Normal-Normal distributions.

¹⁰We could not provide a more general proof because of the complexity of the calculations. However, we conjecture that these results will also hold with more than five firms.

¹¹Please reference the related literature section of this paper for an overview of the value of commitment theory.

strategic complement to the quantity strategy of the firm that moves just before it, it is a strategic substitute to the quantity strategy of all his other preceding movers. For example, the follower firm produces more than the leader firm by strategic complementarity when there are two firms in the market. This finding is consistent with the discussed real-life observations that the second-mover firm achieves a higher market share than the first-mover firm.

We demonstrate four effects that explain our main result about the total welfare rankings between Cournot and Stackelberg mode of conducts¹². The first effect is the traditional first-mover advantage strategic effect. Since followers' production decisions are strategic substitutes of their predecessors' production decisions in the absence of demand signaling through outputs, early-mover firms preempt their followers by investing in a large capacity compared to Cournot. Therefore, the first-mover advantage induces total output and total welfare to increase under the Stackelberg competition (as compared to Cournot).

The second effect is the information acquisition effect. In a leader-followers game, followers perfectly infer the private demand signal of their predecessors by observing their predecessors' output choices in the perfect revealing equilibrium. Therefore, they are better informed about demand compared to a Cournot oligopolist. Accordingly, followers are likely to produce more when the demand is high and produce less when it is low under the Stackelberg competition relative to the Cournot competition. This implies that prices are less responsive to the underlying demand shock under the Stackelberg competition. This greater price stability induces higher welfare, implying this effect favors Stackelberg over Cournot in welfare terms.

The third effect, so called the signaling effect, accounts for the negative output effect of information acquisition by the followers on their predecessors. Under this effect, any non-last mover firm is reluctant to choose a high quantity to avoid signaling high demand to its successor(s). The last mover, on the other hand,

¹²As these effects appear only under the Stackelberg competition, the introduction of them has at least two more consequences. First, they are also useful to understand the profit rankings between firms under the Stackelberg game. Second, consider a Stackelberg game between a manufacturer and two retailers when demand is stochastic. The manufacturer's (i.e., leader's) price can send a signal to the retailers (i.e., followers) about the potential demand. In such a setting, our introduced effects can be reintroduced under a price-setting set-up. For instance, Gal-or et al. (2008, 2011) only show the presence of inference effect in this game. This effect is closely related to the signaling effect as discussed in the text.

does not face any signaling effect because there is no firm following it. This effect, which is also absent under Cournot competition, is one of the main effects that makes simultaneous move games more favourable to society than sequential move games.

The last effect is the negative externalities of information acquisition effect. Since demand shocks are common for all firms, a more informed follower firm lead the residual demand of its competitor(s) less variable. Therefore, the rival firms will have less value from exploiting their demand information. This lower variability in the demand intercept of the competitor firms would translate into lower total welfare if they have some prior of demand initially as in our set-up. That implies that negative externalities of information acquisition effect favor the Cournot competition over the Stackelberg competition in welfare terms.

In sum, only the first two effects favor the Stackelberg competition over the Cournot competition in terms of welfare. Nevertheless, the impact of the signaling and negative externalities of information acquisition effects dominate the impact of the remaining two effects on total welfare. Accordingly, the simultaneous-move quantity setting game generates more total welfare than its sequential counterpart.

Our results lead to a number of implications. First, there are implications on market concentration and welfare. The traditional view is that the oligopoly power effects of having a higher market concentration¹³ induce less competition in the market and is therefore harmful for the society. Nevertheless, Daughety (1990) and several other authors following him argue that the Stackelberg mode of conduct is both more concentrated and more efficient in total welfare terms than the Cournot mode of conduct. Therefore, concentration is beneficial to the society. But our reversal total welfare ranking between Stackelberg and Cournot competitions suggests that this result would be the opposite when firm's strategies involve learning about demand from the actions of their competitors. As a result, the traditional view is reobtained in this paper.

Second, there are implications on the relationship of prices and average firm profits with market concentration. The often observed empirical result is that the correlation of prices and average firm profits with various measures of concentra-

¹³There might also be cost-efficiency effects of a change in concentration. However, concentration is only generated by the non-cooperative nature of competitive interactions among firms in our set-up. Therefore, those kinds of effects are absent here.

tion is positive.¹⁴ Unfortunately, perfect information about demand models may fail to give intuition about these empirical results. Specifically, the Stackelberg model generates lower average total firm profits and lower prices than does the Cournot model under perfect information about demand. Hence, there is a negative correlation of concentration with both average firm profits and prices for some particular domain of problems. Nevertheless, we can give partial explanation to these empirical results by using our incomplete information demand setting as the more concentrated Stackelberg industries create not only higher average firm profit but also higher prices than do the less concentrated Cournot industries.¹⁵

Lastly, there are implications on the impact of market concentration on the merger incentives of firms. We consider a cost-efficient horizontal merger to monopoly in both Cournot and Stackelberg duopoly markets. As both firms merge to monopoly, the highest level of concentration is achieved under both types of markets. Hence, market concentration is likely to be increased more after a merger in less concentrated pre-merger industries (i.e., Cournot). In addition, if the Cournot-Nash mode generates higher pre-merger welfare than the Stackelberg mode of conduct (as in our model, but not under perfect information), then a lower level of efficiency gains would be sufficient to allow mergers in the Stackelberg markets than in the Cournot markets. That also partially explains why the U.S. Federal Trade Commission (FTC) directs its resources toward mergers made among firms that most increased market concentration¹⁶ as it requires more efficiency gains to allow mergers. Altogether, mergers are more likely in more concentrated (i.e., less competitive) industries.¹⁷

We also argue that there is a discontinuity between the Stackelberg equilibrium of the perfect information game and the limit of Stackelberg perfect revealing equilibria of the incomplete information games as the noise of the demand information vanishes to zero.¹⁸ For intuition, consider a two-firm set-up. There are

¹⁴Visit Weiss (1974), Sherer (1980), and many others (will be added) for profit-concentration studies. For a survey of around 25 articles about price-concentration studies between 1989-2004, we refer to Newmark (2004).

¹⁵A more complete analysis would be comparing prices, profits, and welfare with concentration indexes among different asymmetric market structures. We will be focusing on this important policy question in a separate research paper.

¹⁶This observation is based on a January 2013 report of FTC about its horizontal merger investigations from 1996 to 2011.

¹⁷See Hackbarth and Miao (2012) for a similar finding.

¹⁸This discontinuity existed in Gal-Or (1987), Shinkai (2000), and Gal-Or et al. (2008, 2011).

two counter effects, namely the signaling effect and the first-mover advantages. First, as the leader has more information about demand, the information value of the additional observation about demand is more significant to the follower. Therefore, the follower firm relies more on the leaders' selected output in updating his beliefs about demand. But when the leader's signal gets very precise, this dependency is very high and therefore the leader loses its preemptive capability through the indirect revelation of information through output observation by its follower. The signaling effect is the highest in such a case and leads the leader to not produce. Second, as the followers' signal gets also very precise, the first-mover advantages is also the highest and the leader is willing to produce the perfect information outcome. Altogether, these two counter effects lead the leader to produce a level lower than the perfect information level as both firms' precisions get very precise. Hence, we have the observed discontinuity. However, we also argue that if the leader is sufficiently uninformed about demand as compared to the follower and firms have access samples from the same distribution, then the first-mover advantages can still dominate the signaling effect.

In the last section of the paper, we study the robustness of our results to the product differentiation. We show that our main results are mostly valid in a duopoly model, where firms produce imperfectly substitutable products.

This article is organized as follows. In Section 2, we survey some related literature. In Section 3, we provide the set-up. In Sections 4 and 5, we state the Cournot and Stackelberg quantity setting oligopoly models respectively and derive equilibrium outputs. Section 6 compares the Stackelberg model with the Cournot model according to price, total output, total profit, consumer surplus and total welfare. Concluding remarks including extensions and policy implications of the analysis follow in Section 7. Section 8 concludes. Proofs are in the Appendix.

However, they did not point it out. For example, in the latter two companion papers, as the noise of demand information of the manufacturer's and two retailers' signals converge to zero (i.e., as $s_0, s_1, s_2 \rightarrow 0$), the limit of manufacturer's perfect revealing equilibrium expected wholesale price under the no-sharing of information and one-sided information sharing regimes respectively go to $E[p^w] = aE[y^{NS}] \rightarrow \frac{a(2b+d)}{8b-2d}$ and $E[p^w] = aE[y^{PSi}] \rightarrow \frac{a(6b^2-bd+d^2)}{4b(4b-d)}$ by Lemma 1 of Gal-Or et al. (2008). But when the products are differentiated, i.e., $0 < d < b$, both of these limit equilibrium prices are different than the perfect information wholesale price of the manufacturer, namely $a/2$. This argument shows that there is, indeed, such a discontinuity.

2 Related Literature

Over the last three decades, there has been a growing research literature that compares Cournot and Stackelberg equilibrium outcomes under perfect, imperfect, and incomplete information settings. Nevertheless, demand is known in all of these papers. To start with, Boyer and Moreaux (1986) argue that both total output and welfare are higher while total profits are lower in the Stackelberg competition than in the Cournot competition when demand is linear and costs are constant in an n -firm set-up. Later, Boyer and Moreaux (1987) show that these rankings are robust to the product differentiation. Both Anderson and Engers (1992) and Okuguchi (1999) conclude that the same result holds when the demand function has a more general structure. Robson (1990) allow for decreasing average costs (by considering fixed costs) and show that if a Cournot equilibrium has three or more firms, it is necessarily less efficient than the Stackelberg equilibrium, in a total surplus sense. However, when the number of firms is two, the rankings are unambiguous. Albaek (1990) shows that the Stackelberg competition creates more welfare as compared to the Cournot competition in a duopoly market where firms face cost uncertainty. However, as firms are not interested in the competitor's costs *per se*, the Stackelberg game being considered is not a cost signaling game.

Several researchers also studied two-stage generalized Stackelberg models wherein m leaders and $n - m$ followers, where $m \in [0, n]$, compete. In such a model, $m = 0$ or $m = n$ corresponds to the Cournot model. When $m = 1$ and $n = 2$, the set-up becomes the usual one leader-one follower Stackelberg game. When there is full information about demand, Daughety (1990) finds that the above rankings between Cournot and Stackelberg competitions are preserved under linear demand and symmetric constant marginal costs. In that regard, concentration is beneficial to the society. When this model is generalized by assuming general demand and (symmetric) cost structures, Ino and Matsumura (2012) shows that when the number of followers is endogenized by considering the free entry of followers, beneficial concentration always occur and the Stackelberg game generates more welfare than the Cournot game. Later, Ino and Matsumura (2013) consider the same general demand and cost structures, but rather they exogenize entry. They demonstrate that if m is sufficiently high, the rankings are preserved. However, if m is low enough, the Cournot model might create higher welfare than the Stack-

elberg model with linear demand and quadratic costs. The intuition is that when marginal costs are increasing, total production costs are minimized when all firms produce the same output level in the Cournot model.

There is also a rich literature about stochastic Stackelberg games, where m leaders and $n - m$ followers compete and only the leader firm faces demand uncertainties. However, a comparison of Cournot and Stackelberg competitions is not presented in these papers. For instance, De Wolf and Smeers (1997, $m = 1$, n is general), Xu (2005, $m = 1$, n is general), and De Miguel and Xu (2009, m and n are general) compute and characterize the Stackelberg equilibrium in such set-ups by using general demand and cost structures. Liu (2005, $m = 1$, $n = 1$), however, investigates whether there are first- or second-mover advantages when both demand and costs are linear.

Some researchers also explore relationships between commitment and observability in Stackelberg theory literature. Those would include Bagwell (1995), Van Damme and Huckers (1997), Güth et. al (1998), Maggi (1999), Várdy (2004), and Morgan and Várdy (2007). For example, Bagwell (1995) considers noisy-leader games where there is uncertainty about the leader's discrete action. He shows that the set of pure-strategy Nash equilibrium outcomes for the Stackelberg game coincides exactly with the set of pure-strategy equilibrium outcomes for the associated Cournot game. Hence commitment (i.e., first-mover advantage) may have no value if there is (even a slight) noise in the observation of the leader's action. Similarly, Maggi (1999) considers two types of uncertainty faced by the follower about both the leader's action and the leader's type (i.e., cost uncertainty). He shows that as the ratio of the noise about the leader's action to the noise about cost uncertainty goes to zero (infinity, respectively), the leader's output approaches the Stackelberg output (the simultaneous move output). Hence the value of commitment under private information is restored for low noise levels.

To the best of our knowledge, our paper is the first to compare Cournot and Stackelberg models by introducing a welfare analysis when there is demand uncertainty. It is well known from Gal-Or (1987), Mailath (1993), Shinkai (2000), and Gal-Or et al. (2008, 2011) that the first-mover advantage is reduced due to signaling distortions when there is demand uncertainty under a two-firm (three-firm in the latter three) Stackelberg setting.¹⁹ But the Cournot competition is not

¹⁹Similar kinds of signaling distortions are also discussed in Spence (1973) and Milgrom and

presented in these papers. Moreover, we show most of our results in an n -firm setting, where the Stackelberg competition is challenging to study as one can realize from the above market settings. Our paper also contributes to the value of commitment literature by presenting the discontinuity argument in the case of demand uncertainty. In particular, first-mover advantages decrease even for a slight noise in the observation of demand in the perfect revealing equilibrium.

3 Set-Up

We use similar notation as Gal-Or (1987) and Shinkai (2000). Let $N = \{1, 2, \dots, n\}$ be a finite set of firms. We consider an oligopolistic market where $n \geq 2$ firms sell a homogeneous good (or service) at a price of p and compete in quantities. Each firm i produces at a production level of q_i . Let $Q = \sum_{i \in N} q_i$ be the aggregate output of production in the market.

There is a continuum of identical consumers with quadratic utility function $U(q_1, q_2, \dots, q_n) + q_0$, where q_0 is the quantity of the numeraire good. Imagine there is a representative consumer. She maximizes consumer surplus:

$$\max_{q_i} CS = U(q_1, q_2, \dots, q_n) - pQ = (a - \mu + u)Q - \frac{bQ^2}{2} - pQ, \quad i = 1, 2, \dots, n \quad (1)$$

where $a > \mu > 0$, u is a random variable with mean μ and variance σ , a is the observed market demand parameter by all firms, and $b > 0$ is the slope of the demand curve.

The maximization problem in (1) with respect to $q_i \geq 0$ gives that the market demand is linear and stochastic of the form²⁰, ²¹:

$$p = a - \mu + u - bQ \quad (2)$$

Each firm faces an identical technology and exhibits constant returns to scale.

Roberts (1982).

²⁰Having the constant term $-\mu$ in the demand function will later lead to notational simplifications. We could have equivalently considered a combined constant term, e.g., A , which is equal to $a - \mu$.

²¹It is common in oligopoly literature with incomplete information to assume linear demand. For example, see Vives (1984), Li (1985), Gal-Or (1987), Raju and Roy (2000), Gal-Or et al. (2008, 2011), and Vives (2011).

We normalize the unit cost of production to zero.²² These symmetry assumptions about firms' technologies will later ensure that each firm is active in the market. Profit of each firm is expressed as π_i and defined by $\pi_i = pq_i$ under zero cost normalization. Each firm is risk-neutral and maximizes its expected profits.

No firm can observe the realized value of the prior random variable u , but each firm i can observe the realized value of its own private signal y_i on u . Firms have access to samples from the same distribution. Therefore, each private signal is the sum of true demand and additional "white noise," namely $y_i = u + e_i$. Assume that conditional on u , y_1, y_2, \dots, y_n are independent and identically distributed random variables and

$$E(y_i|u) = u, \quad Var(y_i|u) = m, \quad i = 1, 2, \dots, n \quad (3)$$

Hence each firm observes a private signal, which is an unbiased estimator of the true demand. The precision of each signal is symmetric and given by $1/m$. Whereas a signal is uninformative as $m \rightarrow \infty$, it is perfectly informative when $m = 0$ and we return to the full information case. Hence, a lower (higher, respectively) m means that all firms are more (less) informed about the market demand at the same magnitude.

Note that $E(y_i) = E(E(y_i|u)) = E(u) = \mu$ by the law of iterated expectations and $E(u^2) = Var(u) + (E(u))^2 = \sigma + \mu^2$. Together with (3), we get

$$\begin{aligned} E(y_i^2|u) &= Var(y_i|u) + (E(y_i|u))^2 = m + u^2 \\ Var(y_i) &= E(y_i^2) - (E(y_i))^2 = E(E(y_i^2|u)) - \mu^2 = E(m + u^2) - \mu^2 = \sigma + m \\ E(y_i y_j) &= E(E(y_i y_j|u)) = E(E(y_i|u)E(y_j|u)) = E(u^2) = \sigma + \mu^2, \quad i \neq j \\ E(uy_i) &= E(E(uy_i|u)) = E(uE(y_i|u)) = E(u^2) = \sigma + \mu^2 \end{aligned} \quad (4)$$

The following linearity and symmetry assumptions about the posterior expected values in an n -firm oligopoly generalize the respective assumptions of Shinkai (2000) in a three-firm oligopoly.

Assumption 1. *Let (i_1, i_2, \dots, i_n) be an order on the set of firms N .*

²²With positive unitary costs, $a - \mu$ should be interpreted as the difference between the deterministic intercept of the demand and the unit cost.

i) For all $j \in N$, there exist constants $\alpha_{j0}, \alpha_{j1} \in \mathbb{R}$ such that

$$E(u|y_{i_1}, y_{i_2}, \dots, y_{i_j}) = \alpha_{j0} + \alpha_{j1}(y_{i_1} + y_{i_2} + \dots + y_{i_j}) \quad (5)$$

ii) For all $k \in N$, all $l \in N \setminus \{1, 2, \dots, k\}$, there exist constants $\beta_{j0}, \beta_{j1} \in \mathbb{R}$ such that

$$E(y_{i_l}|y_{i_1}, y_{i_2}, \dots, y_{i_k}) = \beta_{k0} + \beta_{k1}(y_{i_1} + y_{i_2} + \dots + y_{i_k}) \quad (6)$$

Since the variance term m is a common parameter for all firms, then it is natural to make the symmetry assumption.²³ Specifically, the constant terms in (5) and (6) depend only on the number of signals in the expectations but not on the identity of them.

The linearity assumption is crucial for the derivation of our results and is frequently assumed in the oligopoly literature with private information.²⁴ Assumption 1 together with the linear demand and constant marginal costs assumptions will ensure that the equilibrium quantities of the Cournot and Stackelberg models, which will be formally defined in the next section, are linear in their arguments. Linear equilibria are tractable, particularly in the presence of private information, have desirable properties like simplicity, and have proved to be very useful as a basis for empirical analysis (Vives, 2011). The prior-posterior distribution pairs that satisfy Assumption 1 include the Gamma-Poisson, Beta-Binomial, and Normal-Normal distributions (DeGroot (1970), Gal-Or (1987), Shinkai (2000)). Since we wish to impose non-negativity constraints on the intercept of the demand function, the most appropriate distributions are the first two, where both u and each y_i ($i = 1, 2, \dots, n$) have a positive support. The coefficients of the posterior expectations given in Assumption 1 are derived in the following lemma.

Lemma 1. *Let $m, \sigma \in \mathbb{R}_+$ and Assumption 1 hold. Let also (i_1, i_2, \dots, i_n) be an order on the set of firms N .*

i) For all $j \in N$,

$$E(u|y_{i_1}, y_{i_2}, \dots, y_{i_j}) = \frac{m\mu}{m + j\sigma} + \frac{\sigma}{m + j\sigma}(y_{i_1} + y_{i_2} + \dots + y_{i_j}) \quad (7)$$

²³This symmetry assumption cannot be valid when each firm i 's signal has a different precision, i.e., $1/m_i \neq 1/m_j$ for $i \neq j$. See Gal-Or (1987) for more discussion.

²⁴For instance, the linearity assumption is also assumed by Vives (1984), Li (1985), Gal-Or (1987), Raju and Roy (2000), Gal-Or et al. (2008, 2011).

ii) For all $k \in N$, all $l \in N \setminus \{1, 2, \dots, k\}$,

$$E(y_i | y_{i_1}, y_{i_2}, \dots, y_{i_k}) = \frac{m\mu}{m + k\sigma} + \frac{\sigma}{m + k\sigma}(y_{i_1} + y_{i_2} + \dots + y_{i_k}) \quad (8)$$

Proof: See the Appendix.

Ex-ante expected total welfare is expressed as the sum of consumer surplus, which is specified by (1), and producer surplus (pQ):

$$E[TW] = E[U(q_1, q_2, \dots, q_n)] = E[(a - \mu + u)Q - \frac{bQ^2}{2}] \quad (9)$$

For example, when price is zero, total quantity would be $(a - \mu + u)/b$. Therefore, expected ex-ante total welfare is $E[TW] = (a^2 + \sigma)/(2b)$ by (4) and (9).

Next, we present Cournot and Stackelberg oligopoly models, after which we compare both types of equilibria and derive the conclusions.

4 The Cournot Oligopoly Model

In the Cournot game, firms simultaneously set quantities after privately observing their signals. A Bayesian equilibrium of the Cournot game is that for all $k \in N$, it holds that $q_k \in \operatorname{argmax}_x E[\pi_k(x, \mathbf{q}_{-k}) | y_k]$, where we let \mathbf{q}_{-k} be the vector of quantities produced by all firms other than k . We next derive the equilibrium quantities.

Theorem 1. *The unique Bayesian equilibrium of the Cournot game is $(q_{1,C}^*(N), q_{2,C}^*(N), \dots, q_{n,C}^*(N))$, where*

$$q_{i,C}^*(N) = \frac{a}{b(n+1)} + \frac{\sigma(y_i - \mu)}{b(2m + \sigma(n+1))}, \quad i = 1, 2, \dots, n \quad (10)$$

where C denotes the Cournot competition.

Proof: See the Appendix.

We also show that the expected Cournot output always equal the Cournot certainty output. In the perfect information case, $m = 0$ and $y_i = \mu$. In such a case, (10) simplifies to $q_{i,C}^*(N) = \frac{a}{b(n+1)}$. In addition, since $E(y_i) = \mu$,

$E(q_{i,C}^*(N)) = \frac{a}{b(n+1)}$, as desired. Having said that, total expected production in the Cournot game is therefore given by

$$E(Q_C^*(N)) = \frac{na}{b(n+1)} \quad (11)$$

In what follows, we calculate expected Stackelberg equilibrium values in a stochastic demand environment and then compare total expected equilibrium outputs under both types of competition.

5 The Stackelberg Oligopoly Model

We assume that each firm chooses its output level after observing the private signal but before realizing the actual demand in a hierarchical Stackelberg quantity setting oligopoly game. Without loss of any generality, suppose firms choose outputs sequentially in the order of their firm numberings. In that regard, firm one, being the Stackelberg leader, first chooses its output quantity, then firm two (a follower) does, then firm three does and so on. Firms are assumed to pre-commit to their production of outputs. Let R_+ and Y_i denote the pure strategy space and firm i 's private signal's strategy space respectively. Firm one chooses its optimal quantity of output after observing its private signal y_1 . Its strategy is denoted by $F_1(y_1)$ where $F_1 : Y_1 \rightarrow R_+$. Firm two can condition its quantity of output on both its private signal y_2 and on the output quantity q_1 chosen by the leader. Hence, its optimal strategy is denoted by $F_2(y_2, q_1)$ where $F_2 : Y_2 \times R_+ \rightarrow R_+$. In general, firm k , $k > 1$, being the $(k-1)^{th}$ follower, conditions its quantity of output on its private signal y_k and the output quantities q_1, q_2, \dots, q_{k-1} chosen by the previous firms. Accordingly, the optimal strategy (or say the quantity decision rule) followed by the $(k-1)^{th}$ follower is denoted by $F_k(y_k, q_1, q_2, \dots, q_{k-1})$ where $F_k : Y_k \times R_+ \times \dots \times R_+ \rightarrow R_+$.

As also discussed in Gal-Or (1987) and Shinkai (2000), there might be two possible kinds of equilibria, 1) Perfect revealing equilibria and 2) Partially revealing equilibria.²⁵ In the first one, the follower firm k can always invert the

²⁵Tirole (1995, Chapter 11, pp. 450-453) argues that when the state of market demand can be two types (e.g., high and low) and $n = 2$, among all separating and pooling equilibria, the only kind of equilibria that survives intuitive criterion in the Stackelberg model is the separating (perfect revealing) equilibria. Assume now that $n = 2$ and only the follower firm faces demand

functions F_1, F_2, \dots, F_{k-1} and perfectly infers the private information of its predecessors. Gal-Or (1987) further demonstrates the existence and uniqueness of perfect revealing equilibrium for a wide range of parameter values in a duopoly set-up. The uniqueness of equilibrium is desirable for answering policy relevant questions. In the second type of equilibria, the leader's decision rule is mostly discontinuous and includes flat regions. But if it is continuous and bounded, then partially revealing equilibria may arise only at the boundaries of permitted quantities of output. Moreover, there are infinitely many partially revealing equilibria, which creates further equilibrium selection problems. Besides these arguments, we want to compare Cournot-Bayesian equilibrium with the Stackelberg equilibrium. Since the Cournot equilibrium is continuous and linear in the firm's signals, it is reasonable to study a continuous and linear equilibrium in the Stackelberg case. Lastly, several researchers such as Gal-Or (1987), Shinkai (2000), and Gal-Or et al. (2008, 2011) studied perfect revealing equilibrium in this literature to answer many policy relevant questions. This type of equilibrium turns out to be a very useful tool to provide a link between real life observations and theory. We would also like to extend the first two paper's analyses to an n -firm oligopoly setting. The bottom line is that many economists desire the separating equilibrium for a variety of reasons (for example, intuition about the result) and often many of the refinements cooperate with this objective. In the light of these arguments, it is more appealing for us to study perfect revealing equilibria.

DEFINITION. A strategy combination $(q_1^*(N), q_2^*(N), \dots, q_n^*(N))$ is a *Stackelberg perfect revealing equilibrium* if it satisfies the following n -system of equations:

uncertainties. Under this assumption, Mailath (1993) shows that there is a unique perfect revealing equilibrium and this equilibrium is the only type of equilibrium that survives $D1$ criterion in the Stackelberg model, where the state of the demand can take three discrete types. Moreover, both Mailath (1989) and Janssen and Maasland (1997) show that the same result also holds when the state of market demand is continuum of types (as in our paper). We conjecture that the perfect revealing equilibrium is also the only type of equilibrium that survives $D1$ criterion in our model but proving this claim requires an involved analyses and therefore, we leave it as an open question for a separate research paper.

private signals are assumed to be independent of the number of firms in the market. Accordingly assume that for each $i \in S$, firm i 's Stackelberg equilibrium strategy is linear of the form

$$q_{i,SQ}^*(S) = F_i^S(y_i, q_1, q_2, \dots, q_{i-1}) = \gamma_{i0s} + \gamma_{i1s}q_1 + \gamma_{i2s}q_2 \dots + \gamma_{i,i-1,s}q_{i-1} + \gamma_{iis}y_i \quad (12)$$

where for $k < i$, γ_{jks} denotes firm i 's output reaction to the changes in the production level of firm k , q_k , ceteris paribus; $\gamma_{iis} > 0$ is firm i 's own production sensitivity to the changes in its private signal; and $s = |S|$ denotes the coefficient identity for the game that is considered. Hence in an n -firm problem, there are $n(n+3)/2$ unknown coefficients and it is very difficult to find them in an efficient manner especially when n is high.²⁷

Since we are particularly interested in perfect revealing equilibrium, the inverse functions of $F_1^{S_1}(y_1), F_2^{S_1}(y_2, q_1), \dots, F_{s-1}^{S_1}(y_{s-1}, q_1, q_2, \dots, q_{s-2})$ exist by the definition of equilibrium and are linear by (12). Hence the information set that each player has depends on both her private signal and its predecessors' signals inferred from their output observations. Consequently firm i 's information vector is an i -dimensional vector of the form²⁸

$$\mathbf{y}_i = (y_i, y_1 = F_1^{S_1^{-1}}(q_1), y_2 = F_2^{S_1^{-1}}(q_1, q_2), \dots, y_{i-1} = F_{i-1}^{S_1^{-1}}(q_1, q_2, \dots, q_{i-1})) \quad (13)$$

The way we find the equilibrium is more constructive than that of Gal-Or (1987) and that of Shinkai (2000). Let $n \geq 3$ and consider any $s \in \{2, 3, \dots, n-1\}$. We consider two Stackelberg games played among firms in markets $S_1 = \{1, 2, \dots, s\}$ and $S_2 = \{1, 2, \dots, s+1\}$. In both pre-entry and post-entry markets, firms move according to their number orderings and firms in S_1 are assumed observe the same private signal. We proceed in two steps. First, we will show that the expected profit maximization problems of each firm $i \in \{1, 2, \dots, s-1\}$ in S_1 and S_2 markets are constant multiples of each other. Thus, its best reply remains the same after the entry. The main intuition is that the residual demand left to firms in $S_1 \setminus \{s\}$ remains the same following the entry. Because of the iterative nature of

²⁷Even three-firm calculations require Mathematica calculations several pages long to derive the equilibrium coefficients. See Shinkai (2000) for more discussion.

²⁸Here for each $j = \{1, 2, \dots, s-1\}$, $F_j^{S_1}(\cdot)$ is inverted with respect to q_j . Since the y_i 's are independent of the number of firms in the market, so is \mathbf{y}_i .

the problem, each such firm acts as if it were a monopolist facing the residual demand curve inherited from the preceding movers. Hence, each such firm's output is independent of the number of firms that follow it in the hierarchy. This observation is similar to the one in the full information case where the same equivalence is also present for firm s for a wide class of demand functions (See Anderson and Engers (1992)).

Second, we show that firm s 's optimal quantity, $q_{s,SQ}^*(S_1)$, decreases to $q_{s,SQ}^*(S_2) = \frac{\sigma}{m+\sigma(s+1)}q_{s,SQ}^*(S_1)$ following the entry of firm $s+1$ into the market as a last mover. These two findings are then sufficient for us to derive the equilibrium quantities of the original game with N firms. All of these claims will be derived in the proof of the next section's theorem.

5.2 The Stackelberg Equilibrium Quantities

The Stackelberg equilibrium quantities are derived in the next theorem²⁹:

Theorem 2. *Let $n \geq 2$ and $m, \sigma \in \mathbb{R}_+$. A unique linear perfect-revealing equilibrium $(q_{1,SQ}^*(N), q_{2,SQ}^*(N), \dots, q_{n,SQ}^*(N))$ exists in the n -player Stackelberg quantity competition game. These equilibrium quantity strategies (best responses) are given by*

$$q_{1,SQ}^*(N) = F_1(y_1) = \frac{a\sigma m + \sigma^2(a - \mu) + \sigma^2 y_1}{2b(\sigma + m)(2\sigma + m)},$$

²⁹One can easily show that when $n = 2$ or $n = 3$, then the equilibrium quantities of firms are positive for any positive realization of the signals, y_1, y_2 , and y_3 . However, if $n \geq 4$, then the quantity of any follower of firm three might be negative for some extreme positive realizations of signals and other parameters of the model. Firms are constrained to choose positive quantities. For convenience and analytical tractability, we can ignore this and get negative quantities for certain combinations of a and the signals. The probability of such an event can be made arbitrarily small by appropriately choosing the variances of the model (e.g., see Vives (1984), pp. 77, Vives (1999), Chapter 8, Raju and Roy (2000)).

For each $i \in \{2, 3, \dots, n-1\}$,

$$\begin{aligned} q_{i,SQ}^*(N) &= F_i(y_i, q_1, q_2, \dots, q_{i-1}) \\ &= \frac{b(i\sigma + m)((2i-3)\sigma + 2m)q_{i-1}(N) - b\sigma^2 \sum_{j \leq i-2} q_j(N) + \sigma^2(a - \mu) + \sigma^2 y_i}{2b(i\sigma + m)((i+1)\sigma + m)}, \end{aligned}$$

$$\begin{aligned} q_{n,SQ}^*(N) &= F_n(y_n, q_1, q_2, \dots, q_{n-1}) \\ &= \frac{b(n\sigma + m)((2n-3)\sigma + 2m)q_{n-1}(N) - b\sigma^2 \sum_{j \leq n-2} q_j(N) + \sigma^2(a - \mu) + \sigma^2 y_n}{2b\sigma(n\sigma + m)} \end{aligned}$$

Equilibrium quantities are iteratively calculated as $q_{1,SQ}^*(N) = F_1(y_1)$ and for each $i \in N \setminus \{1\}$, $q_{i,SQ}^*(N) = F_i(y_i, q_{1,SQ}^*(N), q_{2,SQ}^*(N), \dots, q_{i-1,SQ}^*(N))$.³⁰

Proof: See the Appendix.

Note that when $n = 2$, this theorem is a special case of Gal-Or (1987)'s result.³¹ On the other hand, Shinkai's (2000) result is a special case of this theorem when $n = 3$ and $b = 1$.³² Gal-Or (1987) further shows the uniqueness of perfect revealing equilibrium in the two-firm case.³³

5.3 Strategic Substitutes versus Strategic Complements

We say that quantity strategy q_i is a strategic substitute (or complement respectively) to quantity strategy q_j , $i \neq j$, if the best response of firm i to an increase in the quantity of firm j is to decrease (increase) its production (equivalently if $\frac{\partial^2 \pi_i}{\partial q_i \partial q_j} < 0$ (> 0)). In this section, we investigate strategic substitutability versus

³⁰In the proof of Lemma 5(ii), which is stated in the Appendix, we provide expected equilibrium quantities of firms, which are all positive for all $a, m, \sigma > 0$.

³¹Gal-Or (1987)'s set-up coincides with ours when we let $h = \sigma$, and $m_1 = m_2 = m$ in her paper. Note that the constant deterministic demand parameter a in that paper is changed to $a - \mu$ in this paper to simplify the notations.

³²Shinkai (2000) did not check whether the second order conditions are satisfied or not in his proof. Our proof of Theorem 2 further completes his proof by showing that the second order conditions are, indeed, fulfilled.

³³Theorem 2 shows that there is a unique linear perfect revealing equilibrium. We also conjecture that there is a unique perfect revealing equilibrium that is linear with more than two firms.

complementarity relationships among firms' strategies in our Stackelberg setting. One can use the best responses provided in Theorem 2 to find these relationships. Lemma 2 summarizes our results.

Lemma 2. *Let $m, \sigma \in \mathbb{R}_+$.*

i) for each $k \in N \setminus \{1, n\}$ and $l = k - 1$, $\frac{\partial q_{k,SQ}^}{\partial q_l} = \frac{2m+(2k-3)\sigma}{2(m+\sigma(k+1))} > 0$,*

ii) $\frac{\partial q_{n,SQ}^}{\partial q_{n-1}} = \frac{2m+(2n-3)\sigma}{2\sigma} > 0$,*

*iii) for each $n \geq 4$, each $k \in N \setminus \{1, 2, n\}$, each $l \in \{1, 2, \dots, k - 2\}$,
, $-\frac{1}{2} < \frac{\partial q_{k,SQ}^}{\partial q_l} = -\frac{\sigma^2}{2(m+k\sigma)(m+\sigma(k+1))} < 0$,**

iv) for each $n \geq 3$ and $l \in N \setminus \{n - 1, n\}$, $-\frac{1}{2} < \frac{\partial q_{n,SQ}^}{\partial q_l} = -\frac{\sigma}{2(m+n\sigma)} < 0$.*

These derivations show that although the quantity of any follower firm is a strategic complement to the quantity of the one that moves just before it (parts *i* and *ii*), it is a strategic substitute to all other preceding movers to that follower firm (parts *iii* and *iv*). In the following figure, we depict these findings with four firms.

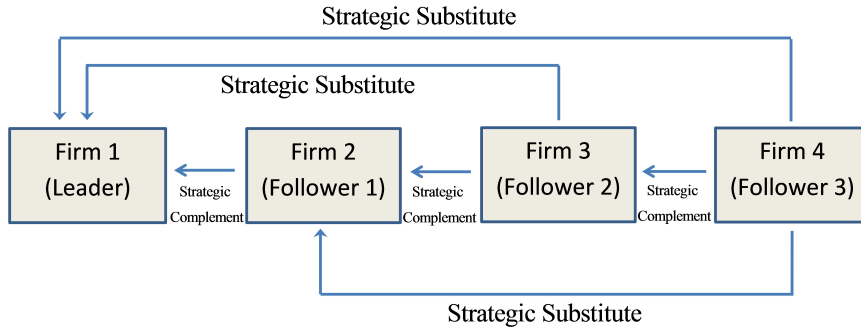


Figure 1: **Strategic Effects:** In this figure, we provide the strategic relationships among quantity strategies of four firms in the Stackelberg setting based on Lemma 2. Whereas the quantity of any follower is a strategic complement to his immediate predecessor, it is a strategic substitute to the one of all his other preceding movers.

In order to provide our intuition for the observed relationships, we introduce the direct revelation benchmark game, where firms directly reveal their private

information to their followers instead of letting them infer it from their observable quantity of output.³⁴ In this game, there is no signaling of information through output observations as in our Stackelberg model. Therefore, a comparison of the equilibrium quantities of Stackelberg and direct revelation games will show us that how information acquisition affects the followers' optimal decisions.

In a direct revelation game, let the equilibrium strategies of each firm $i \in N$ be linear of the form

$$\begin{aligned} q_{i,D}^*(N) &= H_i(y_1, y_2, \dots, y_i, q_1, q_2, \dots, q_{i-1}) \\ &= \xi_0 + \xi_1 y_1 + \xi_2 y_2 + \dots + \xi_i y_i + \xi_{i+1} q_1 + \dots + \xi_{2i-1} q_{i-1} \end{aligned} \quad (14)$$

where subscript D denotes direct revelation. In Lemma 3, we derive the equilibrium strategies and equilibrium payoffs of the game with direct revelation.

Lemma 3. *A unique linear equilibrium $(q_{1,D}^*(N), q_{2,D}^*(N), \dots, q_{n,D}^*(N))$ exists in the direct revelation game. These equilibrium strategies are given by, for each $i \in N$,*

$$q_{i,D}^*(N) = H_i(y_1, \dots, y_i, q_1, \dots, q_{i-1}) = \frac{a - \mu - b \sum_{j \in \{1, 2, \dots, i-1\}} q_j + E[u|\mathbf{y}_i]}{2b} \quad (15)$$

Equilibrium quantities are iteratively calculated as $q_{1,D}^(N) = H_i(y_1)$ and for each $i \in N \setminus \{1\}$, $q_{i,D}^*(N) = H_i(y_1, y_2, \dots, y_n, q_{1,D}^*(N), q_{2,D}^*(N), \dots, q_{n,D}^*(N))$.*

Proof: The proof is very similar to the proof of Theorem 2 and is therefore skipped.

Observe by Lemma 3 that the quantity strategies of any follower is a strategic substitute to the quantity strategy of its predecessors, i.e., $\frac{\partial q_{i,D}^*}{\partial q_j} = \frac{-1}{2}$, $i > j$. Accordingly comparing this observation with the findings in Lemma 2 shows that there exists a *positive conjectural variation effect* that is beneficial to the follower firms in our Stackelberg model. This effect increases all slope parameters in the best response functions of followers compared to a benchmark level of $-1/2$. From the perspective of the last-mover, it does not really matter whether the information

³⁴This benchmark is not really implementable because it requires that the follower is capable of verifying the information transmitted by the leader. If a mechanism of verification does not exist, then there might be an incentive for early-mover firms to report untruthfully signals lower than their true realizations (Gal-Or et al., 2008). Thus, the followers generally suspect the accuracy of the signaling information received from their predecessors (Gal-Or, 1987).

is directly or indirectly revealed. However, indirect inferences about demand will decrease non-last-movers' preemptive abilities (i.e., first-mover advantages). It is, therefore, this positive effect that might convert strategic substitutability relations into strategic complementarity ones.

Consider now an increase in any non-last mover firm's, say firm j , output level associated with a positive demand shock. This increase initiates chain reactions in the optimal output strategies of the followers of firm j . Firm $j + 1$ is the first to respond to this increase by further increasing its output level through its strategic complementarity relationship with firm j . But an increase in both firm j 's and $j + 1$'s quantities induce prices to decrease too much if firms from $j + 2$ up to n do not react. They, however, find it optimal to make a negative adjustment by decreasing their output levels. Consequently, there will be a moderate increase in the market price. If, otherwise, all followers of firm j increase their quantities following this positive demand shock, then price will be too low, which will create incentives for them to deviate by reducing their productions.

6 Stackelberg versus Cournot Oligopoly

6.1 Total Output and Price Comparisons

As prices are negatively correlated with total output, it is sufficient to show the ranking of total equilibrium output between Cournot and Stackelberg competitions. In that regard, define $\Delta E(Q^*(N))$ as the difference of total expected equilibrium production in the Cournot market from the one in the Stackelberg market, i.e., $\Delta E(Q^*(N)) = E(Q_C^*(N)) - E(Q_{SQ}^*(N))$. Our next theorem establishes that the simultaneous move game lead to higher total expected output and lower expected price than the sequential move game irrespective of how noisy both the demand shocks and firms' private signals are.

Theorem 3. *Let $n \geq 2$ and $m, \sigma \in \mathbb{R}_+$. Whereas total expected equilibrium market output is higher, the expected equilibrium price is lower in the Cournot game than in the Stackelberg game. (i.e., $\Delta E(Q^*(N)) > 0$)*

Proof: See the Appendix.

We identify two effects that induce our result about output rankings between Cournot and Stackelberg equilibrium outcomes. The first effect is the traditional *first-mover advantage effect*. Recall that the firms' optimal decisions involve strategic substitutability relationships in the absence of demand signaling through outputs by Lemma 2. For this reason, each firm (save the last mover) selects a high output in order to induce subsequent movers to cut back. Consequently, the first-mover advantage induces total output to increase under the Stackelberg competition (as compared to Cournot).

The second effect is the *signaling effect*³⁵, which is absent under the Cournot competition. As noted in the last section, followers become more responsive to the output changes of their preceding movers under the positive conjectural variation effect. That creates an incentive to invest in a lower capacity for the firms that signal their demand information. Therefore, under the signaling effect, any non-last mover firm would like to produce a low quantity to signal its followers that the demand is low. In that regard, this negative output effect favors the Cournot competition over the Stackelberg competition. In sum, there is one effect favoring Stackelberg over Cournot and one effect favoring Cournot over Stackelberg. It turns out that total negative output effects due to signaling information dominate total positive output effects due to moving early. Thus, the total expected equilibrium output is unambiguously higher in the Cournot case.

We finally consider the impact of these effects on the individual output levels of firms when $n = 2$. Our results are summarized in Lemma 4.

Lemma 4. *Let $n = 2$ and $m, \sigma \in \mathbb{R}_+$.*

$$E[q_{2,SQ}^*(\{1, 2\})] > E[q_C^*(\{1, 2\})] > E[q_{1,SQ}^*(\{1, 2\})]$$

Proof: See the Appendix.

The leader firm faces both signaling and first-mover advantages effects unlike a Cournot duopolist. Remark by Lemma 2 that the reaction function of the follower is upwards sloping. Therefore, the leader firm's output reduction incentive, which

³⁵In a set-up where two retailers infer demand information from the price chosen by one manufacturer., Gal-Or et al. (2008, 2011) identify the presence of an inference effect, which is the cousin of signaling effect in price setting games. Specifically, because of this inference effect, the manufacturer would like to set a low wholesale price to signal to the retailer that the demand is low.

is induced by the signaling effect, is likely to dominate its output expansion incentive, which is induced by the first-mover advantages. Consequently, it produces less than a Cournot duopolist in expectation. Nevertheless, since the follower firm does not face any of above effects, it can increase its output aggressively through strategic complementarity. Indeed, it produces more than a Cournot duopolist in expected terms.

6.2 Profit, Consumer- and Total Surplus Comparisons

In this section, we provide our main results. We compare total expected equilibrium profit, consumer surplus and total surplus between the Stackelberg and Cournot competitions. We start with the total expected profit comparisons. Total expected equilibrium profit is defined as

$$E[\Pi^*(N)] = E[p^*Q^*(N)] = E[(a - \mu + u - bQ^*(N))Q^*(N)] \quad (16)$$

Now, let the total expected profit difference between these models be denoted as $\Delta E(\Pi^*(N)) = E(\Pi_C^*(N)) - E(\Pi_{SQ}^*(N)) = nE(\pi_C^*(N)) - \sum_{i \in N} E(\pi_{i,SQ}^*(N))$.

Theorem 4. *Let $n \in \{2, 3, 4, 5\}$ and $m, \sigma \in \mathbb{R}_+$. Total expected profit is higher in the Stackelberg game than in the Cournot game.³⁶ (i.e., $\Delta E(\Pi^*(N)) < 0$)*

Proof: See the Appendix.

Before providing the main effects that generate the total profit rankings in this theorem, we derive consumer surplus and total welfare rankings between our models. Ex-ante total expected equilibrium welfare can be deduced by summing up consumer and producer surplus, which are respectively defined in (1) and (16), as:

$$E[TW^*(N)] = E[(a - \mu + u)Q^*(N) - \frac{b(Q^*(N))^2}{2}] \quad (17)$$

³⁶Proving this claim in a general n -firm framework requires tedious algebraic calculations. Given this constraint, we prove the claims up to a 5-firm set-up and conjecture that the claims hold with more than five firms.

Based on this definition, let $\Delta E(TW^*(N))$ be the difference of the Cournot competition total equilibrium welfare from the Stackelberg competition total equilibrium welfare, i.e., $\Delta E(TW^*(N)) = E(TW_C^*(N)) - E(TW_{SQ}^*(N))$. Similarly, the same difference for the consumer surpluses is defined as $\Delta E(CS^*(N)) = E(CS_C^*(N)) - E(CS_{SQ}^*(N))$. Next, we present a main result.

Theorem 5. *Let $n \in \{2, 3, 4, 5\}$ and $m, \sigma \in \mathbb{R}_+$. Both consumer and total expected equilibrium welfare are higher under the Cournot competition than under the Stackelberg competition.³⁷ (i.e., $\Delta E(TW^*(N)) > 0$ and $\Delta E(CS^*(N)) > 0$.)*

Proof: See the Appendix.

As previously discussed, first-mover advantages increase early-mover firms' production incentives to capture a bigger pie of the market. This increase in output levels translates into not only higher consumer surplus and total welfare but also lower total profits as noted under perfect information models that compare Stackelberg and Cournot competitions. By analogy, the adverse consequences of the signaling effect on output investment incentives of the early-mover firms is expected to generate lower consumer surplus and total welfare; and higher total profits. As a result, the domination of the signaling effect over the first-mover advantages effect partially explains the observed rankings between two types of competition in Theorem 5.

Apart from the above two opposite output effects, we finally point out two welfare effects in order to completely justify the rankings in Theorem 5. These two effects are called as *the information acquisition effect* and *negative externalities of information acquisition effect*. They do not account for total output differences, but rather they give the possibility that even when the expected outputs are the same between two types of competition, total profit, consumer surplus, and total welfare rankings are likely to be different. Specifically, the followers are better informed than their predecessors in a leader-followers game. Therefore, they are likely to produce more when the demand is high and produce less when it is low under the Stackelberg competition compared to the Cournot competition. This implies that prices are less responsive to the underlying demand shock under the

³⁷See Footnote 36.

Stackelberg competition. This greater price stability (and higher output production volatility) induces higher welfare, implying that the information acquisition effect favors Stackelberg over Cournot in terms of total profits, consumer surplus and total welfare.³⁸

Information acquisition of a follower firm also imposes a negative externality to its competitor(s) in the following sense. Since demand shocks are common for all firms, a more informed follower firm lead the residual demand of its competitor(s) less variable. This lower variability in the demand intercept of the competitor firms would translate into lower total welfare if they have some prior of demand initially as in our set-up. Therefore, this effect is likely to decrease total welfare. That implies that negative externalities of information acquisition effect favors the Cournot competition over the Stackelberg competition in welfare terms.

Among these four effects, the signaling and negative externalizes of information acquisition effects are the dominant ones. Accordingly, the Cournot mode of conduct is socially desirable compared to the Stackelberg mode of conduct from both consumer and total welfare point of views. This result sharply diverges from traditional efficiency rankings between these two game settings.

We now study two simple examples in order to better understand these two welfare effects. In the first one, we let each firm a monopoly for the good it produces in a duopoly set-up. Since demand intercepts are common for both firms, the follower firm still learns the signal of the leader firm in a perfect revealing equilibrium. Hence this type of information acquisition affects the welfare analysis between Cournot and Stackelberg competitions. In the second example, we provide more intuition about the welfare effects of the negative externalities of information acquisition of firms on their competitors.

Example 1: Let $N = \{1, 2\}$ and $p_i = a - \mu + u - bq_i$, $i = 1, 2$, where $b > 0$ is the slope of demand. We follow all other assumptions of our original model. The Stackelberg perfect revealing equilibrium outputs of firms will be derived in the extension Section 7.5 as

$$q_{1,DS}^* = \frac{a}{2b} + \frac{\sigma(y_1 - \mu)}{2b(m + \sigma)} \quad \text{and} \quad q_{2,DS}^* = \frac{a}{2b} + \frac{\sigma(y_1 + y_2 - 2\mu)}{2b(m + 2\sigma)} \quad (18)$$

³⁸The intuition is similar to the welfare effects of third degree price discrimination, where third degree price discrimination (lower price stability) reduces welfare when the expected quantities are the same.

where DS refers to the differentiated Stackelberg game. Similarly, Cournot-Bayesian equilibrium outputs of firms are given as

$$q_{i,DC}^* = \frac{a}{2b} + \frac{\sigma(y_i - \mu)}{2b(m + \sigma)} \quad i = 1, 2 \quad (19)$$

where DC corresponds to the differentiated Cournot game. Note that from the perspective of the first firm, it does not make any difference whether the type of competition is Cournot or Stackelberg. It simply produces the monopoly level of output in both set-ups. Therefore, $q_{1,DS}^* = q_{1,DC}^*$. However, since the signal of the leader is revealed to the follower under the Stackelberg competition, the follower is more informed about demand as the compared to the Cournot competition.³⁹ Comparing (18) and (19) shows that the follower is able to produce more when the observed signal of the leader is greater than $\frac{m\mu + \sigma y_2}{m + \sigma}$ as compared to a Cournot duopolist. For example, consider $y_2 = \mu$. Firm two produces $a/2b$ under the Cournot competition. However, the Stackelberg follower produces more than $a/2b$, when the observed signal of the leader is greater than the average (*i.e.*, $y_1 > \mu$). That implies that there is additional consumer surplus, producer surplus and total surplus generated by producing more in the high states of demand. The opposite is true when the leader's realization of the state of demand is lower than $\frac{m\mu + \sigma y_2}{m + \sigma}$. Nevertheless, the gain in higher states of demand in welfare grounds overwhelms the loss in lower states of demand with a linear demand curve.⁴⁰ Therefore, although the expected output of a firm under both mode of conducts are the same, expected- total profits, consumer surplus, and total welfare are all higher under sequential move games. To put the arguments formally, we have

$$\begin{aligned} E[TW_{DS}^*] - E[TW_{DC}^*] &= 3(E[CS_{DS}^*] - E[CS_{DC}^*]) = 3/2(E[\Pi_{DS}^*] - E[\Pi_{DC}^*]) = \\ &= \frac{3m\sigma^2}{8b(m + \sigma)(m + 2\sigma)} > 0 \end{aligned} \quad (20)$$

³⁹Despite of its usefulness, this example might not be very realistic at first glance. When the goods are totally differentiated, one might think that the follower firm does not care about the demand information revealed by the leader. However, the production decision of the leader firm can still send a signal to the follower firm about the general macroeconomic conditions in an economy, which can affect the common demand parameter among firms.

⁴⁰Kühn and Vives (1995) provide an excellent graphical analysis in a two discrete state of demand set-up to explain why providing more information about demand to a monopoly firm induces higher welfare in quantity-setting games under linear demand structures.

This example is useful in two ways. First, it shows the pure effect of information acquisition on welfare grounds. Since the follower's decision rule is independent of the leader's action and the follower's decision does not affect the residual demand function of the leader, all other three effects are absent here. Second, when the degree of differentiation is high enough in the market, both the consumer surplus and total welfare are no longer higher in simultaneous move-games by continuity of the parameter values. In other words, some of our main results of the paper will not be robust when firms produce almost monopoly products. Nevertheless, we will see in Section 7.5 that our main results hold for a very wide range of parameter values in a differentiated good environment.

Example 2: We finally give the intuition for the discussed welfare effects of the negative externalities of information acquisition of the follower firm on the leader firm in our original set-up when $n = 2$. We define the residual demand function that the leader firm faces as $R_1(q_2) = a - \mu + u - bq_2$ and plot the expected residual demand by the red dotted line in Figure 1. We also draw the representative residual demands in high and low states as compared to the expected demand curve, which are respectively denoted by R_1^H and R_1^L . Now consider that the follower is more informed about demand. Accordingly, the residual demand function of the leader becomes less volatile and both R_1^H and R_1^L will lie more closer to the expected residual demand curve. This can be seen through noting that firms have interactions through having common demand intercepts and producing substitutable products. Intuitively, the acquisition of information by the follower makes the leader react more actively in terms of output to the demand shock. This gives the leader firm less value from exploiting the information it has about demand. As a result, the leader firm charges a lower price when the demand is higher than the expected residual demand.⁴¹ That creates a loss in the leader firm's expected profit of an amount equivalent to Area $B + D$ in Figure 1. However, consumers are not affected from this change because the linearity of the demand curve ensures that the loss in consumer's surplus (Area B) is exactly equal to the gain in consumer surplus (Area C). In sum, expected total welfare is expected to decrease in high states of demand. By analogy, the leader charges a higher price when the demand is lower than the expected residual

⁴¹We also refer to Kühn and Vives (1995, pp. 13-18) for a geometrical representation of similar arguments when the state of demand is two in a monopolistic competition set-up.

demand, which creates a gain in expected total welfare. But the loss in both the leader firm's profits and total welfare in high states of demand is greater than the gain in them in bad states of demand if the leader initially has some information about demand.⁴² Altogether, having less variable residual demand is likely to decrease not only the leader's profits but also total welfare while leaving consumer surplus unchanged.

6.3 Discontinuity of Equilibria in the Noise of the Signals

As $m \rightarrow 0$, the precision of the signals converges to infinity and therefore we approach the perfect information case. It is straightforward to see that the Stackelberg oligopoly expected equilibrium quantities in a world of perfect information coincide with the limit of direct revelation equilibrium quantities, which are given by (15), as $m \rightarrow 0$.⁴³ In particular, perfect information expected equilibrium quantities can be calculated by letting $m = 0$ and $y_i = \mu$, $i = 1, 2, \dots, n$ in (15) as

$$E[\mathbf{q}_{PQ}^*] = (E[q_1^*], E[q_2^*], \dots, E[q_n^*]) = \left(\frac{a}{2b}, \frac{a}{4b}, \dots, \frac{a}{2^i b}, \dots, \frac{a}{2^n b}\right) \quad (21)$$

where PQ refers to the perfect information quantity-setting game. However, as $m \rightarrow 0$, expected Stackelberg equilibrium quantities in the indeterministic case do not converge to (21):

$$\lim_{m \rightarrow 0} E[\mathbf{q}_{SQ}^*] = \left(\frac{a}{4b}, \frac{a}{8b}, \dots, \frac{a(2i-1)!!}{2^i(i+1)!b} \mid_{i < n}, \dots, \frac{a(2n-3)!!}{2^{n-1}(n)!b}, \frac{a(2n-1)!!}{2^n(n)!b}\right) \quad (22)$$

from Lemma 5(ii) (See Appendix), where $(2n-1)!! = \prod_{i \in N} (2n-1)$. Immediate comparisons between (21) and (22) show that there is a discontinuity between the Stackelberg equilibrium of the perfect information game and the limit of perfect revealing equilibria of the incomplete information games as the noise of the demand information vanishes to zero. Indeed, this kind of divergence might lead to some unexpected results.

We extend our model by considering firm specific accuracies of signals in order

⁴²If the leader has no prior about demand, the gain in the leader's profit in bad states is exactly equal to the loss in its profit in good states.

⁴³The corresponding linear demand functional form in the deterministic case is also given by $p = a - \mu + u - Q$. Although the demand is still stochastic, firms know exactly which state of the demand they are in. The perfect information output levels of firms under the Stackelberg competition is provided in Vives (1988).

to provide our intuition for this discontinuity result in the two-firm case. As in Gal-Or (1987), let $1/m_i$, $i = 1, 2$, denotes the precision of the signal of firm i . Under this generalization, the unique pure strategy separating equilibrium of the game can be derived by letting $h = \sigma$, $a = a - \mu$, and $\theta = \mu$ in Proposition 1 of Gal-Or (1987) as^{44, 45}

$$\begin{aligned} q_{1,SQ}^{**}(\{1, 2\}) &= \frac{m_1\sigma((a - \mu)(\sigma + m_1) + m_1\mu + \sigma y_1)}{2b\Delta(\sigma + m_1)} \\ q_{2,SQ}^{**}(\{1, 2\}) &= \frac{m_1\sigma(a - \mu + y_2)}{2b\Delta} + q_1\left(-\frac{1}{2} + \frac{m_2(\sigma + m_1)}{m_1\sigma}\right) \end{aligned} \quad (23)$$

where $\Delta = \sigma(m_1 + m_2) + m_1m_2$. Expected quantities are:

$$E[q_{1,SQ}^{**}(\{1, 2\})] = \frac{a\sigma m_1}{2b\Delta} \quad (24)$$

$$E[q_{2,SQ}^{**}(\{1, 2\})] = \frac{a\sigma m_1}{2b\Delta} \left(\frac{1}{2} + \frac{m_2(m_1 + \sigma)}{m_1 + \sigma}\right) \quad (25)$$

We subsequently consider two extreme cases, where perfect revealing equilibrium is not formally defined. We, indeed, make the arguments in ϵ neighbourhoods of the selected m_1 and m_2 parameter values.

As an extreme case, when $m_2 = 0$, the follower does not use the quantity of output selected by the leader as a source of information about the demand. But then both the leader and the follower choose the expected quantity of output that would have been selected in a world of perfect information. Indeed, as m_2 goes to 0, $\frac{\partial q_{2,SQ}^{**}}{\partial q_1}$ converges to $-\frac{1}{2}$ by (23). Hence, positive conjectural variation effect vanishes. In such a case, there is not any signaling effect and first-mover advantages are the highest (i.e., $\lim_{m_2 \rightarrow 0} E[q_{1,SQ}^{**}] = a/(2b)$ and $\lim_{m_2 \rightarrow 0} E[q_{2,SQ}^{**}] = a/(4b)$).

The other extreme case is when $m_1 = 0$.⁴⁶ In this case, the follower relies heavily on the observation of q_1 in updating his beliefs about the demand. That imposes an extra constraint on the leader's decision rule. As $\lim_{m_1 \rightarrow 0} \frac{\partial q_{2,SQ}^{**}}{\partial q_1} = \infty$ from (23), the follower becomes very responsive to the changes in the production of

⁴⁴ $h = \sigma$ refers to the case where each firm has access samples from the same distribution as in our model. Further, a is replaced with $a - \mu$ and $\theta = \mu$ in our paper.

⁴⁵When $m_1 = m_2 = m$, these equilibrium quantities coincide with Theorem 2 for $n = 2$.

⁴⁶For a similar discussion, see paragraph 5 at pp.288 of Gal-Or (1987).

the leader. Accordingly, the leader is forced to contract its output so as to induce reduced production by the follower. As a result, the leader loses its preemptive capability and $E[q_{1,SQ}^{**}]$ becomes zero. This is the case where the signaling effect has the highest pressure on the leader's production decision.

Now consider any identical pair of precision of signals $(\tilde{m}_1, \tilde{m}_2)$ with $\tilde{m}_1 = \tilde{m}_2 = m$. As m converges to zero, $\frac{\partial q_{2,SQ}^{**}}{\partial q_1}$ converges to $1/2$ by (23). Hence, the positive conjectural variation effect will be bigger than $-1/2$. In such a case, although the first-mover advantage leads the leader to produce $a/(2b)$ (in expected terms), the jump down (signaling) effect forces the leader to not produce based on the above discussion made for the extreme cases. In sum, the leader ends up producing at an intermediate level, which is $a/(4b)$. That induces the follower to produce $3a/(8b)$ in expected terms rather than the perfect information outcome $a/(4b)$. Hence we get the observed discontinuity as m approaches to zero and the signaling distortions play an important role in this nontrivial result. We conjecture that a similar argument to above can be made in an n -firm case and therefore we skip it.

Based on the above discussion, the study herein does not have a conclusion that perfect information outcomes are never achievable as a limit of Stackelberg perfect revealing equilibria. Note to see that for each $m_1 \in \mathbb{R}_+$, there always exists a positive $m_2^* = \frac{m_1\sigma}{2(m_1+\sigma)}$, which is smaller than m_1 , such that the first-mover advantage dominates the signaling effect if and only if $m_2 < m_2^*$. This can be seen by observing that $m_2 < m_2^*$ if and only if $\frac{\partial q_{2,SQ}^{**}}{\partial q_1} < 0$ by (23). Thus, if the precision of the follower's signal is sufficiently more informative than the precision of the leader's signal and the firms have access samples from the same distribution, then the Stackelberg duopoly perfect revealing equilibrium strategies might still be close to the Stackelberg perfect information outcomes. Essentially, any firm's expected production is increasing in the precision of its rival's signal (i.e., $\frac{\partial E[q_{i,SQ}^{**}]}{\partial m_j} < 0$ for $i \neq j$ by (24) and (25)). The leader, therefore, prefers to compete with a more informed follower firm to less suffer from the signaling distortions.

The presence of negative signaling distortions also rises the loss in the value of commitment issues in leader-follower games with demand uncertainties. For making our arguments more concrete, consider Mailath's (1993) similar Stackelberg model to us, where only the follower firm faces three discrete types of demand un-

certainties⁴⁷. Typically, the sequencing decision of firms itself depends on a firm's private information. Therefore, in a game with endogenous sequencing, a choice of Cournot sequencing leaves the follower firm with the problem of inferring not only the informed (leader) firm's information but also the quantity the informed firm intends to select. In that regard, the leader firm is not able to avoid the negative signaling distortion by moving simultaneously with the follower because the follower would view a choice of Cournot sequencing as an attempt to hide high demand. Consequently, although the leader may earn lower ex-ante profits than it would earn if it was choosing quantities with the follower (as in our model), it always chooses to move first regardless of its private information in the unique stable (perfect revealing) equilibrium of this described game⁴⁸.

Lastly, Gal-Or argues that if the firms have access to samples from possibly different populations that are partially correlated, then we still have significant divergences from perfect information outcomes for a wide range of parameter values even with asymmetric precision of signals. To be more precise, let u_i be the random variable that affects the demand in the segment observed by firm i . Let $u = (u_1 + u_2)/2$ and $p = a - \mu + u - Q$. Further, assume that $E[y_i|u_i] = u_i$ and $Var[y_i|u_i] = m_i$. Let h measures the degree of correlation between u_1 and u_2 . Similar linearity assumptions to our paper are assumed on the distributions of u and y_i 's. When $h = \sigma$, firms have access distributions from the same samples as in our original set-up. She shows that if h changes between 0 and $2\sigma/3$, then the reaction function of the follower is upwards sloping at any $m_1, m_2 \in \mathbb{R}_+$. This is because the less the degree of correlation between u_1 and u_2 , the more the signaling effect is felt by the leader. Having said that, our proposed rankings between the Cournot and Stackelberg equilibrium outcomes are also expected to be valid for this very large set of parameter values.⁴⁹

⁴⁷Mailath (1989) shows that all results of Mailath's (1993) would hold in the case of continuum types of demand as in this paper.

⁴⁸For more discussion, we refer to Mailath (1993).

⁴⁹For instance, we first introduce the Cournot competition in Gal-Or's two-firm set-up with demand uncertainty. After some involved calculations, one can show that total expected (unique) Cournot equilibrium quantity is $2a/3b$ as in our baseline Cournot model. Subtracting total expected Stackelberg equilibrium outputs (provided in the proof of Observation 2 of Gal-Or's paper) from total expected Cournot equilibrium quantity yields $\frac{a(h^2 + (m_1 + \sigma)(2\sigma - 3h + 2m_2))}{12b\Delta}$, which is positive at any $h \in (0, 2\sigma/3]$ and $m_1, m_2 \in \mathbb{R}_+$. Accordingly, the equilibrium price is higher under the Stackelberg competition than under the Cournot competition.

7 Discussions and Extensions

7.1 Firm Asymmetries and Concentration

Concentration is mostly thought to be a cause for concern.⁵⁰ This concern is based on an intuition derived mainly from considering the symmetric equilibrium of oligopoly models. As the number of firms in the symmetric equilibrium increases, some measures of welfare rises (e.g., total surplus) and some measures of concentration falls (e.g., The Herfindahl-Hirschman Index (or HHI)). However, the differences in firm sizes can also be a cause of concentration. Daughety argues that the foregoing intuition does not carry over to the asymmetric equilibria which seemingly reflect a more realistic picture of the world. His findings in a world of perfect information about demand suggest that social optimality may involve extensive asymmetries in firm sizes. These contradicting results indicate that the sign of the correlation between concentration indexes and welfare is ambiguous.

In the light of these concerns, we analyze how firm asymmetries affect social optimality in our set-up. The HHI is a measure of the size of firms in relation to the industry and is widely accepted by the anti-trust agencies to be an indicator of the amount of concentration among firms. Typically, the Agencies measure each firm's market share based on annual data. They can also consider projected market shares in the relevant market to take into account firms' future competitive significance⁵¹. Thus, we can slightly extend the definition of HHI to capture the demand uncertainties in market shares. In that regard, let $H = \sum_{i=1}^n (E[s_i])^2$, where $E[s_i] = E[q_i]/E[Q(N)]$ is the expected market share of firm i with $\sum_{i=1}^n E[s_i] = 1$. Substituting $\sum_{i=1}^n E[s_i] = 1$ into $H(\cdot)$ and taking minimization over it gives

$$\min_{E[s_i]} (E[s_1])^2 + (E[s_2])^2 + \dots + (E[s_{n-1}])^2 + (1 - E[s_1] - E[s_2] - \dots - E[s_{n-1}])^2 \quad (26)$$

As $\frac{\partial^2 H}{\partial^2 E[s_i]} = 4 > 0$, $H(\cdot)$ is convex in $E[s_i]$. Using the symmetric first-order conditions from (26), one can find that $H(\cdot)$ is minimized at $(E[s_i])^* = 1/n$ for $i \in N$. Remark that as the timing of producing is different under the Stackelberg model, firms have asymmetric expected market shares. This can be seen analytically from Lemma 5(ii) (see the Appendix). Therefore, on average, the Stackelberg model

⁵⁰This paragraph is mostly adopted from Daughety (1990).

⁵¹See 2010 Horizontal Merger Guidelines for more discussion about the calculation of the HHI.

yields a higher HHI than does the Cournot model, where firms behave symmetrically. In other words, the market is, on average, more concentrated under the Stackelberg mode of conduct than under the Cournot mode of conduct. In what follows, we relate our main findings with concentration.

7.2 Welfare and Concentration

Daughety (1990) argues that concentration measures (such as HHI) provide little insight about welfare. His findings reflect that increases in such measures sometimes reflect increases in welfare and sometimes reflect decreases in welfare. For example, as we switch from Cournot to Stackelberg model, both the HHI index and total welfare increase.⁵² In that regard, asymmetry, thus concentration, is beneficial from society's viewpoint. This observation contradicts to the standard IO paradigm that a higher market concentration leads to less competition in the market and therefore is harmful for the society. Nevertheless, this result does not take into account the possibility that early moving firms' quantity strategies might send demand signals to the late-movers in a stochastic demand environment. In such a set-up, total welfare ranking between Stackelberg and Cournot competitions are reversed in an n -firm oligopoly model ($n \leq 5$) irrespective of how noisy both the demand shocks and private demand signals of firms are by Theorem 5. This result suggests that the traditional IO paradigm is likely to be preserved when firm's strategies involve learning about demand from the actions of their competitors. Therefore, HHI provides significant amount of insight about welfare. A policy maker would then prefer the less concentrated Cournot industries over the more concentrated Stackelberg industries if he/she has an impact on the determination of market's mode of conduct.

7.3 Profits, Prices and Concentration

A similar implication holds for the relationships of prices and profits with concentration. In study after study, a positive correlation of prices and average firm

⁵²It has already been stated in the related literature section that Daughety studies a two-stage model, where m leaders and $n - m$ followers compete. Apart from the incomplete information about demand extension, our model coincides with his model at $m = 0$ or $m = 2$ (Cournot) and $m = 1$ (Stackelberg) when $n = 2$. In the future, we are planning to study Daughety's model in our incomplete information setting to better assess the link between concentration and welfare.

profits with various measures of concentration has been found as we have already noted in the introduction.⁵³ Unfortunately, perfect information about demand models may fail to support these empirical results. Specifically, a higher concentration, which is caused by switching from the Cournot model to the Stackelberg model, is likely to be associated with both lower average firm profits and lower prices. But remark by Theorems (3) and (4) that the more concentrated Stackelberg competition generates both higher total profits and higher prices than does the less concentrated Cournot competition. Therefore, our foregoing analysis in a world of incomplete information can partially support these empirical results.

7.4 Merger Incentives and Concentration

As a last application, we perform a merger analysis to answer an important policy question: Are mergers easier to be allowed under more concentrated or less concentrated industries? We consider a cost-efficient horizontal merger to monopoly in both Cournot and Stackelberg duopoly markets.⁵⁴ Our previous set-up and assumptions about the parameters of the model is still valid. However, we consider firms that have symmetric marginal costs, i.e., c , in the pre-merger case. Accordingly, the deterministic demand parameter a can be considered as net of the marginal cost level of firms, i.e., $a = \bar{a} - c$. Following a merger, we assume that there are efficiency gains and the merged firm's marginal cost level becomes c_e , where $c_e \in [0, c)$. Accordingly, let $\hat{a} = \bar{a} - c + c_e$. Our welfare analysis and comparisons between simultaneous and sequential move settings performed in the main text corresponds to the pre-merger case. We now let two firms merge to a monopoly horizontally. Under both Cournot and Stackelberg competitions, the merged firm maximizes the expected joint profits of firms one and two.

$$\max_{q_1, q_2} E[\pi_M] = E[(\hat{a} - \mu + u - b(q_1 + q_2))(q_1 + q_2) | y_1, y_2] \quad (27)$$

⁵³Price-concentration studies are widely thought to overcome potential problems on profit-concentration ratios. For example, prices are much more easily observed than economics profits. Further, they are not subject to the many accounting conventions that complicate studying profits.

⁵⁴Since our main concern of the paper is not to explain the relationship between merger incentives of firms and market concentration, we just wanted to give the main intuition by studying a simple horizontal merger model. One can come up with a more sophisticated merger model that does not necessarily lead mergers to monopoly.

where π_M denotes the total profits of the merged firm. Since the merged firm's problem is symmetric between its facilities, then FOC's yield equilibrium quantities as

$$q_1^* = q_2^* = \frac{\hat{a} - \mu + E[u|y_1, y_2]}{4b} \quad (28)$$

Based on (4) and (28), the post-merger expected welfare under both types of competitions reduces to

$$TW_M^* = \frac{3\hat{a}^2(m + 2\sigma) + 6\sigma^2}{8b(m + 2\sigma)} \quad (29)$$

We assume that a merger is allowed if it generates sufficiently large efficiency gains to increase total welfare.

Theorem 6. *Let $n = 2$. The minimum level of efficiency gains to allow mergers to monopoly is lower under the Stackelberg competition than under the Cournot competition.*

Proof: See the Appendix.

Based on this lemma, anti-trust authorities can be more suspicious about mergers in symmetric industries. The main intuition is easy to see. Since the post-merger welfare is the same under both Cournot and Stackelberg competitions, then the pre-merger welfare is the determinant factor in answering our policy question. But since the pre-merger welfare is lower under the Stackelberg competition than under the Cournot competition (but not under perfect information about demand), then it would be reasonable to require less efficiency gains to allow mergers in more concentrated Stackelberg industries. In addition, the highest level of post-merger concentration is achieved under both types of markets as both firms merge to monopoly. Hence, market concentration is likely to be increased more after a merger in less concentrated pre-merger industries (i.e., Cournot). That also partially explains why FTC directs its resources toward horizontal mergers that are made among firms that most increased market concentration as it requires less efficiency gains to allow mergers. Altogether, mergers are more likely in more concentrated industries.

7.5 Differentiated Products

We have assumed that firms produce homogeneous products until this section. We would like to discuss the robustness of our results in an environment where two firms produce heterogeneous products. Let there be two firms competing in quantities. Each firm i ($=1,2$) produces a differentiated product at a price level of p_i and at a production level of q_i . The representative consumer maximizes consumer surplus

$$\max_{q_i} CS = (a - \mu + u)(q_1 + q_2) - \frac{b(q_1^2 + q_2^2)}{2} - b\lambda q_1 q_2 - \sum_{j=1}^2 p_j q_j \quad (30)$$

where $a > \mu > 0$, u is a random variable with mean μ and variance σ , a is the observed market demand parameter by all firms, b is the slope of the demand, and $\lambda \in [0, 1]$ is an inverse measure of product differentiation in the market. When $\lambda = 1$, products are perfect substitutes and no longer differentiable as in the main text. On the other hand, when $\lambda = 0$, products are unrelated and each firm is a monopoly for the product it produces. The maximization problem in (30) yields a linear demand curve with a stochastic term in the intercept:

$$p_i = a - \mu + u - bq_i - b\lambda q_j, \quad i \neq j, \quad i, j = 1, 2 \quad (31)$$

We normalize the unit cost of production to zero.

No firm can observe the realized value of the prior random variable u , but each firm i can observe the realized value of its own private signal y_i on u . Let both u and each y_i ($i = 1, 2, \dots, n$) have a full positive support. we refer to Section 3 for further assumptions on the prior u and private signals y_1 and y_2 .⁵⁵

We now present both Cournot and Stackelberg duopoly models with heterogeneous products. Later, we compare both models.

7.5.1 Differentiated Cournot Duopoly Model

We first assume that two firms simultaneously decide their quantities after getting their private signals about demand. Since firms cannot draw inferences about the private signals of their competitors, each firm's equilibrium quantity strategy only depends on its own signal.

⁵⁵We apologize for this economic way of presenting this model.

Theorem 7. Let $n = 2$ and $m, \sigma \in \mathbb{R}_+$. The unique Bayesian equilibrium of the differentiated Cournot game is $(q_{1,DC}^*, q_{2,DC}^*)$, where

$$q_{i,DC}^* = \frac{a}{b(2 + \lambda)} + \frac{\sigma(y_i - \mu)}{b(2m + \sigma(2 + \lambda))}, \quad i = 1, 2 \quad (32)$$

Proof: See the Appendix.

When $\lambda = 1$, the competition among firms is fierce as goods are perfect substitutes of each other as we assumed in the baseline model. Note also that the Cournot equilibrium quantities in Theorem 1 at $n = 2$ coincide with the ones in Theorem 7 at $\lambda = 1$. On the other extreme, when firms are monopolies for the good they produce ($\lambda = 0$), the expected monopoly quantity is $a/2$ which is the same with the one in the perfect information case. This observation is consistent with our findings under quantity competition stated in Section 4.

7.5.2 Differentiated Stackelberg Duopoly Model

In the Stackelberg model, we assume that each firm chooses its quantity level after observing the private signal but before realizing the actual demand. Without loss of any generality, suppose firm one, being the Stackelberg leader, first chooses its quantity level, then does firm two (follower) after observing the quantity level chosen by the leader. Let R_+ and Y_i denote the pure strategy space and firm i 's private signal's strategy space respectively. Firm one chooses its optimal quantity level after observing its private signal y_1 . Its strategy is denoted by $G_1(y_1)$ where $G_1 : Y_1 \rightarrow R_+$. Firm two can condition its optimal quantity level on both its private signal y_2 and on the quantity level q_1 chosen by the leader. Hence, its optimal strategy is denoted by $G_2(y_2, p_1)$ where $G_2 : Y_2 \times R_+ \rightarrow R_+$.

As before, we focus on perfect revealing equilibria, which can be defined in a similar fashion to the one in the main text. Formally,

DEFINITION. A strategy combination $(q_{1,DS}^*, q_{2,DS}^*)$ is a *Stackelberg perfect revealing equilibrium* if it satisfies the following two-system of equations:

$$\begin{aligned}
& \forall y_1 \in Y_1, \\
& \quad q_{1,DS}^* = G_1(y_1) = \operatorname{argmax}_{q_1 \in R_+} E[\pi_1(q_1, G_2(y_2, q_1), u) | y_1] \\
& \forall y_2 \in Y_2, \forall G_1(y_1) = q_1 \in R_+, \\
& \quad q_{2,DS}^* = G_2(y_2, q_1) = \operatorname{argmax}_{q_2 \in R_+} E[\pi_2(q_1, q_2, u) | y_2, q_1]
\end{aligned} \tag{33}$$

where DS refers to the differentiated Stackelberg quantity competition.

In order to derive the Stackelberg perfect revealing equilibrium, we conjecture that firms have linear equilibrium quantity strategies. Accordingly, let $q_{1,DS}^* = \alpha_0 + \alpha_1 y_1$ and $q_{2,DS}^* = \beta_0 + \beta_1 y_2 + \beta_3 q_1$ for some constants $\alpha_0, \alpha_1, \beta_0, \beta_1, \beta_3 \in \mathbb{R}_+$. In the proof of Theorem 8, we derive these constants and obtain the equilibrium quantities.

Theorem 8. *Let $n = 2$, $m, \sigma \in \mathbb{R}_+$, and $\lambda \in [0, 1]$. Under Assumption 1, a unique linear perfect revealing equilibrium $(q_{1,DS}^*, q_{2,DS}^*)$ exists in the two-player differentiated Stackelberg (DS) quantity competition game. These equilibrium quantity strategies (best responses) are given by*

$$q_{1,DS}^* = G_1(y_1) = \frac{(2m(1-\lambda) + \sigma(4-3\lambda))(am + (a-\mu+y_1)\sigma)}{2b(m+\sigma)(m+2\sigma)(2-\lambda^2)},$$

$$q_{2,DS}^* = G_2(y_2, q_1) = \frac{(a-\mu+y_2)\sigma}{2b(m+2\sigma)} + q_1 \left(\frac{(2-\lambda^2)(m+\sigma)}{2(1-\lambda)m + (4-3\lambda)\sigma} - \frac{\lambda}{2} \right)$$

Equilibrium quantities are iteratively calculated as $q_{1,DS}^ = G_1(y_1)$ and $q_{2,DS}^* = G_2(y_2, q_{1,DS}^*)$.*

Proof: See the Appendix.

As $a > \mu$ by assumption, both $q_{1,DS}^*$ and $q_{2,DS}^*$ are positive for any positive realizations, namely, y_1 and y_2 . Observe also by Theorem 8 that $\frac{\partial q_{2,DS}^*}{\partial q_1} = \frac{(2-\lambda)(2m+(2-\lambda)\sigma)}{4(1-\lambda)m+2(4-3\lambda)\sigma} > 0$. Therefore, the equilibrium quantity of the follower is a strategic complement to the leader's quantity for any degree of heterogeneity among products.

We are now ready to compare differentiated Cournot and Stackelberg quantity competitions in terms of expected- quantities, total profits, consumer surplus, and

total welfare.

7.5.3 Differentiated Cournot versus Differentiated Stackelberg

Define expected equilibrium total profits as $E[\Pi_D] = E[p_1q_1 + p_2q_2]$ where the subscript D refers to product differentiation. Adding profits to consumer surplus given in (30) and taking an expectation over the resulting equation yields expected total welfare as

$$E[TW_D] = E[(a - \mu + u)(q_1 + q_2) - \frac{b(q_1^2 + q_2^2)}{2} - b\lambda q_1q_2] \quad (34)$$

Now denote the difference between differentiated- Cournot and Stackelberg equilibrium outcomes as $\Delta E[l_D^*] = E[l_{DC}^*] - E[l_{DS}^*]$ for $l \in \{Q, \Pi, CS, TW\}$. The next theorem compares expected- total outputs and total profits of firms between differentiated Stackelberg and Cournot competitions.

Theorem 9. *Let $n = 2$ and $m, \sigma \in \mathbb{R}_+$.*

i) For $\lambda \in (0, 1]$, $\Delta E[Q_D^] > 0$ and for $\lambda = 0$, $\Delta E[Q_D^*] = 0$.*

ii) For $\lambda \in [0, 1]$ and $a \geq \sigma$, we have $\Delta E[\Pi_D^] < 0$.⁵⁶*

Proof: See the Appendix.

Based on Theorem 9, both expected total outputs and total profits are higher under the Cournot competition than under the Stackelberg competition for any degree of substitution between products. This result suggests that output and total profit rankings stated respectively in Theorems 3 and 4 are robust to the product differentiation.

Theorem 10. *Let $n = 2$ and $m, \sigma \in \mathbb{R}_+$.*

i) $\frac{\partial \Delta E[CS_D^]}{\partial a} > 0$ and $\frac{\partial \Delta E[TW_D^*]}{\partial a} > 0$.*

ii) For $a \geq 3\sigma$ and $\lambda \in [0.1, 1]$, we have $\Delta E[CS_D^] > 0$ and $\Delta E[TW_D^*] > 0$.⁵⁷*

⁵⁶In this result, we require the standard deviation of the demand intercept is not too large and smaller than its mean. Shinkai (2000) argues that when the distribution of the demand intercept is approximated by a normal one, the probability of the realized value of it being positive is almost one under the more restrictive assumption that $a \geq 3\sigma$.

⁵⁷We refer to Footnote 56.

Proof: See the Appendix.

We have already concluded in Example 1 of Section 6.2 that when $\lambda = 0$, both rankings stated in Theorem 10(ii) are reversed. However, this theorem shows that for sufficiently high and reasonable deterministic demand parameter a , homogeneous good model rankings of equilibrium welfare and consumer surplus between Cournot and Stackelberg models are preserved for a degree of substitution level greater than 0.1 between firms' products. Hence, even when firms produce differentiated products, the effects of signaling and negative externalities of information acquisition outweighs the effects of information acquisition and first-mover advantages on consumer surplus and total welfare.

8 Conclusion

The purpose of this article is to compare the industry performances of symmetric Cournot and asymmetric Stackelberg mode of conducts in an n -firm oligopoly setting when there is incomplete information about demand. We provided substantial amount of findings that social optimality between these two types of mode of conducts involves extensive symmetry. Moreover, the Stackelberg competition generates higher prices and total profits, yet lower output and consumer surplus than the Cournot competition. These rankings are the opposite to the rankings provided in the previous studies, which were written under perfect information about demand assumption. This divergence drastically changes previous policy recommendations about several real-life case studies.

The industry concentration, which is measured by the HHI, is typically higher in a Stackelberg industry than in a symmetric Cournot industry. In that regard, this study is also a first step towards understanding the relations of concentration (generated by the disparities in firm sizes) to prices, total profits, and welfare in a stochastic demand environment. As the Stackelberg competition creates a lower total welfare than the Cournot competition, concentration is harmful for the society. This harmfulness does not depend on scale economies, marketing advantages, or learning by doing. It derives entirely from the noncooperative nature of firm interactions. Furthermore, the price and profits rankings between two types of mode of conducts support the often observed empirical result that there is a positive correlation of prices and average firm profits with concentration. In the light

of these arguments, perfect information models might not fully provide the realistic picture of the world. We hope that our introduction of private information about demand leads us one step closer to understand the connections between real life observations and theory. Our analysis is also not complete in the sense that one still need to analyze how market concentration changes within different asymmetric mode of conducts.

We determine four effects that explain our results: 1) Information acquisition effect 2) First-mover advantage effect 3) Signaling effect 4) Negative externalities of information acquisition effect. Among these four effects, only the last two effects favor the Cournot competition over the Stackelberg competition in terms of consumer and total welfare. As a consequence of the signaling effect, early-mover firms invest in lower capacity to not signal high demand to their followers in a sequential move setting. As the followers' output reactions to their predecessors' productions increase and sometimes become positive, this negative output effect will typically dominate the positive output effect of traditional first-mover advantages. Further, a follower will have better information compared to a Cournot oligopolist as it learns the private signal of its predecessors in a perfect revealing equilibrium. Having more informed players about demand is expected to generate higher consumer surplus, total profits, and total welfare in a Stackelberg quantity setting game as compared to a Cournot game. Nevertheless, this type of information acquisition by the followers also leads the residual demand function of their competitors less variable as all firms share a common demand intercept. Therefore, the rival firms will have less value from exploiting their demand information. That will translate into not only lower profits but also lower welfare. It turns out that the effects of signaling and negative externalities of information acquisition always dominate the effects of other two on industry performances and we have the observed rankings.

We finally show that there is a discontinuity between the Stackelberg equilibrium of the perfect information game and the limit of Stackelberg perfect revealing equilibria of the incomplete information games as the noise of the demand information vanishes to zero. Hence, the first-mover advantage is reduced at any symmetrical noise level of firms' demand signals. This result contributes to the value of commitment literature in Stackelberg models.

We also challenge the robustness of our results in various dimensions in a

duopoly set-up. Our main model assumes that firms produce homogenous products and the precision of the firms' private signals are symmetric. When we allow for differentiation among firms' products, we see that our main results of the paper hold for a very wide range of parameter values. We also argue that if the follower's signal is very precise as compared to the precision of the leader's signal and the firms have access samples from the same distribution (as in our main model), then the Stackelberg perfect revealing equilibrium outcomes are close to the perfect information equilibrium outcomes. In such a case, the follower firm is reluctant to use the quantity selected by the leader as a source of information about the demand. Therefore, the positive output effect of the first mover advantages is expected to dominate negative output effect of signaling. However, a decrease in the correlation of signals pronounce the signaling effect more. Accordingly, if firms' signals are sufficiently uncorrelated, then the follower's quantity can be still a strategic complement for any degree of asymmetries in firm specific precisions. In sum, one can also expect significant difference between the equilibrium outcomes of perfect and private information about demand environments in such cases.

APPENDIX

Proof of Lemma 1:

Let Assumption 1 hold. Take any $i \in N$. Since the posterior expectations are linear, a result in Erikson (1969) shows that:

$$E(u|y_i) = \frac{z_i}{z_i + R}y_i + \frac{R}{z_i + R}E(u) \quad (.1)$$

where

$$z_i = \frac{1}{E(\text{Var}(y_i|u))} \quad \text{and} \quad R = \frac{1}{\text{Var}(u)} \quad (.2)$$

Since $E(\text{Var}(y_i|u)) = m$ and $\text{Var}(u) = \sigma$, (.1) simplifies to

$$E(u|y_i) = \alpha_{11}y_i + \alpha_{10} = \frac{\sigma}{\sigma + m}y_i + \frac{m\mu}{\sigma + m} \quad (.3)$$

Now, take any $h \in N \setminus \{i\}$. Note that $E(y_h|y_i) = \beta_{10} + \beta_{11}y_i$ by Assumption 1. By using the law of iterated expectations,

$$\begin{aligned} E(y_h y_i) &= E(E(y_h y_i|y_i)) = E(y_i E(y_h|y_i)) = \beta_{10}E(y_i) + \beta_{11}E(y_i^2) \\ E(y_h) &= E(E(y_h|y_i)) = \beta_{10} + \beta_{11}E(y_i) \end{aligned} \quad (.4)$$

Since $E(y_h) = E(y_i) = \mu$, $E(y_h y_i) = \sigma + \mu^2$, and $E(y_i^2) = m + \sigma + \mu^2$ by (4), one can solve the system of equations in (.4) for β_{10} and β_{11} to have

$$E(y_h|y_i) = \beta_{11}y_i + \beta_{10} = \frac{\sigma}{\sigma + m}y_i + \frac{m\mu}{\sigma + m} \quad (.5)$$

Now let (i_1, i_2, \dots, i_n) be an order on the set of firms N . Let $N_r = \{i_1, i_2, \dots, i_r\}$.

i- For any $j \in N$, $E(u|y_{i_1}) = E(E(u|y_{i_1}, y_{i_2}, \dots, y_{i_j})|y_{i_1}) = \alpha_{j0} + \alpha_{j1} \sum_{l \in N_j} E(y_{i_l}|y_{i_1})$ by the law of iterated expectations and Assumption 1(i). Using the symmetry in (.5),

$$E(u|y_{i_1}) = \alpha_{j0} + \alpha_{j1}(y_{i_1} + (j-1)\left(\frac{m\mu}{\sigma + m} + \frac{\sigma}{\sigma + m}y_{i_1}\right)) \quad (.6)$$

Solving (.3) (at $i = i_1$) and (.6) for the constants yields $\alpha_{j0} = \frac{m\mu}{m+j\sigma}$ and $\alpha_{j1} =$

$$\frac{\sigma}{m+j\sigma}.$$

ii- For all $k \in N$, all $l \in N \setminus \{1, 2, \dots, k\}$, $E(y_{i_l}|y_{i_1}) = E(E(y_{i_l}|y_{i_1}, y_{i_2}, \dots, y_{i_k})|y_{i_1}) = \beta_{k0} + \beta_{k1}(y_{i_1} + (j-1)(\frac{m\mu}{\sigma+m} + \frac{\sigma}{\sigma+m}y_{i_1}))$ by the law of iterated expectations and Assumption 1(ii). One can solve this equality and (.5) (at $i = i_1$) for the constants to have $\beta_{k0} = \frac{m\mu}{m+j\sigma}$ and $\beta_{k1} = \frac{\sigma}{m+j\sigma}$ as desired. \square

Proof of Lemma 4: We claim that $E[q_{2,SQ}^*(\{1, 2\})] > E[q_C^*(\{1, 2\})] > E[q_{1,SQ}^*(\{1, 2\})]$ for any $m, \sigma \in \mathbb{R}_+$. Remark that $E[q_C^*(\{1, 2\})] = \frac{a}{3b}$ from Section 4. Using Lemma 5(ii), which will be stated just after the proof of Theorem 2, we have $E[q_{1,SQ}^*(\{1, 2\})] = \frac{a\sigma}{2b(m+2\sigma)}$ and $E[q_{2,SQ}^*(\{1, 2\})] = \frac{a(2m+3\sigma)}{4b(m+2\sigma)}$. Hence $E[q_{2,SQ}^*(\{1, 2\})] - E[q_C^*(\{1, 2\})] = \frac{a(2m+\sigma)}{12b(m+2\sigma)} > 0$ and $E[q_{1,SQ}^*(\{1, 2\})] - E[q_C^*(\{1, 2\})] = -\frac{a(2m+\sigma)}{6b(m+2\sigma)} < 0$ at any $m, \sigma \in \mathbb{R}_+$ as claimed. \square

Proof of Lemma 1:

We first show that if each $k \in N \setminus \{n\}$ uses $q_{k,C}^*(N) = \frac{a}{b(n+1)} + \frac{\sigma(y_k - \mu)}{b(2m+\sigma(n+1))}$, then the unique best response for firm n is to use $q_{n,C}^*(N) = \frac{a}{b(n+1)} + \frac{\sigma(y_n - \mu)}{b(2m+\sigma(n+1))}$. To see this, notice that the expected profit of firm n choosing quantity strategy q_n given other firms strategies and the signal y_n is

$$E[(a - \mu + u - bq_n - b \sum_{k \in N \setminus \{n\}} q_{k,C}^*(N))q_n | y_n] \quad (.7)$$

So the optimal choice of firm n is

$$q_{n,C}^*(N) = \frac{a - \mu + E(u|y_n) - b \sum_{k \in N \setminus \{n\}} E(q_{k,C}^*(N)|y_n)}{2b} \quad (.8)$$

which, after some computations, equals $q_{n,C}^*(N) = \frac{a}{b(n+1)} + \frac{\sigma(y_n - \mu)}{b(2m+\sigma(n+1))}$ by the initial supposition and Lemma 1. Uniqueness follows similarly as in Li (1985). \square

Proof of Theorem 2:

We proceed in four steps to find the equilibrium quantities. To start the analysis, we first consider the pre-entry market.

Step 1) Pre-entry market ($S_1 = \{1, 2, \dots, s\}$)

In this step, we derive firms $s-1$ and s 's best reply functions in the pre-entry market by using backwards induction. We start out by considering the maximization solved by the last follower, firm s . His objective is to:

$$\max_{q_s} E[\pi_s(q_1, \dots, q_s, u)|y_s, q_1, \dots, q_{s-1}] = E[(a - bQ(S_1) + u - \mu)q_s|y_s, q_1, \dots, q_{s-1}] \quad (.9)$$

The first-order condition is:

$$\frac{\partial E[\pi_s(\cdot)|\cdot]}{\partial q_s} = a - bQ_{S_1 \setminus \{s\}}(S_1) - 2bq_s + E(u|y_s, q_1, \dots, q_{s-1}) - \mu = 0 \quad (.10)$$

In the light of (13), (.10) can be rewritten as⁵⁸

$$q_{s,SQ}^*(S_1) = F_1^{S_1}(y_s, q_1, \dots, q_{s-1}) = \frac{a - bQ_{S_1 \setminus \{s\}}(S_1) + E(u|\mathbf{y}_s) - \mu}{2b} \quad (.11)$$

Similarly, we substitute $q_{s,SQ}^*(S_1)$ in firm $s - 1$'s optimization problem to have:

$$\begin{aligned} \max_{q_{s-1}} E[\pi_{s-1}(q_1, \dots, q_{s-1}, q_{s,SQ}^*(S_1), u)|y_{s-1}, q_1, \dots, q_{s-2}] \\ = \frac{1}{2} E[(a - bQ_{S_1 \setminus \{s\}} - E(u|\mathbf{y}_s) + 2u - \mu)q_{s-1}|y_{s-1}, q_1, \dots, q_{s-2}] \end{aligned} \quad (.12)$$

The first-order condition after some manipulations becomes

$$q_{s-1,SQ}^*(S_1) = \frac{a - bQ_{S_1 \setminus \{s,s-1\}}(S_1) - E(E(u|\mathbf{y}_s)|\mathbf{y}_{s-1}) + 2E(u|\mathbf{y}_{s-1}) - \mu}{2b + \frac{\partial E(u|\mathbf{y}_s)}{\partial q_{s-1}}} \quad (.13)$$

where $\frac{\partial E(u|\mathbf{y}_s)}{\partial q_{s-1}} = \frac{\sigma}{(\sigma+m)\gamma_{s-1,s-1,s}}$ by (12), (13), and Lemma 1. Note that $E(E(u|\mathbf{y}_s)|\mathbf{y}_{s-1}) = E(u|\mathbf{y}_{s-1})$ by Lemma 1(i). Altogether, (.13) simplifies to

$$q_{s-1,SQ}^*(S_1) = \frac{a - bQ_{S_1 \setminus \{s,s-1\}}(S_1) + E(u|\mathbf{y}_{s-1}) - \mu}{2b + \frac{\sigma}{(m+\sigma)\gamma_{s-1,s-1,s}}} \quad (.14)$$

But since $\gamma_{s-1,s-1,s}$ is the coefficient in front of y_{s-1} in $q_{s-1}^*(S_1)$ from (12), we have

⁵⁸To illustrate things more clearly, consider $s = 2$. $E(u|\mathbf{y}_2) = \frac{\sigma}{2\sigma+m}(y_2 + F_1^{S_1^{-1}}(q_1)) + \frac{m\mu}{m+2\sigma} = \frac{\sigma}{m+2\sigma}(y_2 + \frac{q_1 - \gamma_{10s}}{\gamma_{11s}}) + \frac{m\mu}{2\sigma+m}$ by Lemma 1(i) and (12). In the text, we sometimes replace $F_1^{S_1^{-1}}(q_1)$ by y_1 , which are technically equal, to simplify the notation. However, we stress that firm two does not observe y_1 directly but it can perfectly infer it by observing firm one's choice of output level, q_1 . In that regard, the partial derivative $\frac{\partial E(u|\mathbf{y}_2)}{\partial q_1}$ is equal to $\frac{\sigma}{(m+2\sigma)\gamma_{11s}}$ rather than zero.

$\gamma_{s-1,s-1,s} = \frac{\sigma/(m+(s-1)\sigma)}{2b+\sigma/((m+s\sigma)\gamma_{s-1,s-1,s})}$ by (.14) and Lemma 1, which can be solved as:

$$\gamma_{s-1,s-1,s} = \frac{\sigma^2}{2b(m + \sigma(s-1))(m + \sigma s)} \quad (.15)$$

Lastly, substituting (.15) into (.14) yields

$$q_{s-1,SQ}^*(S_1) = F_1^{S_1}(y_{s-1}, q_1, \dots, q_{s-2}) = \frac{\sigma(a - bQ_{S_1 \setminus \{s,s-1\}}(S_1) + E(u|\mathbf{y}_{s-1}) - \mu)}{2b(m + s\sigma)} \quad (.16)$$

Step 2) Post-entry market ($S_2 = \{1, 2, \dots, s, s+1\}$)

In this step, we let firm $s+1$ enter into the market and become the last mover. Other firms preserve the order of their moves. In the post-entry market ($S_2 = \{1, 2, \dots, s, s+1\}$), we deduce firm $s-1$'s best reply in addition to the best replies of firms s and $s+1$. In the end, we show that the best responses and therefore the equilibrium output levels of all firms from 1 to $s-1$ are the same between the pre-entry and post-entry markets.

We can find the best responses of firms $s+1$ and s as in the above pre-entry set-up in a symmetric way. Accordingly, replacing s with $s+1$ and S_1 with S_2 respectively in (.11) and (.16) would give

$$q_{s+1,SQ}^*(S_2) = \frac{a - bQ_{S_2 \setminus \{s+1\}}(S_2) + E(u|\mathbf{y}_{s+1}) - \mu}{2b} \quad (.17)$$

and

$$q_{s,SQ}^*(S_2) = \frac{\sigma(a - bQ_{S_2 \setminus \{s,s+1\}}(S_2) + E(u|\mathbf{y}_s) - \mu)}{2b(m + \sigma(s+1))} \quad (.18)$$

Next, we write down the optimization problem of firm $s-1$ as

$$\begin{aligned} & \max_{q_{s-1}} E[\pi_{s-1}(q_1, \dots, q_{s-1}, q_{s,SQ}^*(S_2), q_{s+1,SQ}^*(S_2), u) | y_{s-1}, q_1, \dots, q_{s-2}] \\ & = E[(a - bQ_{S_2 \setminus \{s,s+1\}}(S_2) - bq_{s,SQ}^*(S_2) - bq_{s+1,SQ}^*(S_2) + u - \mu)q_{s-1} | y_{s-1}, q_1, \dots, q_{s-2}] \end{aligned} \quad (.19)$$

Substituting (.17) and (.18) into the above maximization and letting $E(u|\mathbf{y}_{s+1}) = \frac{s\sigma+m}{(s+1)\sigma+m}E(u|\mathbf{y}_s) + \frac{\sigma}{m+(s+1)\sigma}y_{s+1}$ by Lemma 1(i) and rearranging terms yields

$$\max_{q_{s-1}} \frac{1}{2} E[(\theta_s(a - bQ_{S_2 \setminus \{s,s+1\}}(S_2) - \mu - E(u|\mathbf{y}_s)) - \frac{\sigma}{m+(s+1)\sigma}y_{s+1} + 2u)q_{s-1} | \mathbf{y}_{s-1}] \quad (.20)$$

where $0 < \theta_s = \frac{\sigma(2s+1)+2m}{2(\sigma(s+1)+m)} < 1$. Note that $E(u|\mathbf{y}_{s-1}) = E(y_{s+1}|\mathbf{y}_{s-1})$ by Lemma 1 and both u and y_{s+1} are independent of q_{s-1} . In that regard, it is valid to replace y_{s+1} by u in (.20) to have

$$\max_{q_{s-1}} \frac{\theta_s}{2} E[((a - bQ_{S_2 \setminus \{s, s+1\}}(S_2) - \mu - E(u|\mathbf{y}_s) + 2u)q_{s-1} | \mathbf{y}_{s-1})] \quad (.21)$$

Comparing (.12) and (.21) shows that firm $s - 1$'s maximization problem is just multiplied by a constant term θ_s following the entry. That implies that the best reply of firm $s - 1$ in both markets are the same. Now note that all firms face the same price. Accordingly, one can substitute the best reply of firm $s - 1$ into the prices given in (.12) and (.21) to derive firm $s - 2$'s optimization problem in the related markets. As firm $s - 1$'s problems in the pre-entry and post-entry markets are constant multiples of each other, so does firm $s - 2$'s problems in both markets. Therefore, firm $s - 2$'s best reply also remains the same after the entry. The recursive nature of the Stackelberg game eventually ensures that the maximization problem of each firm i ($i \in \{1, 2, \dots, s - 1\}$) in pre-entry market is a constant multiple of its problem in the post-entry market. Consequently, firm i 's best response does not change due to entry and we therefore have

$$q_{i, S_Q}^*(S_1) = q_{i, S_Q}^*(S_2), \quad i = 1, 2, \dots, s - 1 \quad (.22)$$

at the perfect revealing equilibrium. Using (.22), dividing (.18) by (.11) and rearranging terms gives

$$q_{s, S_Q}^*(S_2) = \frac{\sigma}{m + \sigma(s + 1)} q_{s, S_Q}^*(S_1) \quad (.23)$$

Step 3) The Derivation of Equilibrium Quantities

We are now ready to derive the Stackelberg game equilibrium quantities. We analyze three cases.

Case 1: Leader's best reply:

We start from a pre-entry market with firm one and add new firms one by one as last movers until reaching N . But we have $q_{1, S_Q}^*(N) = q_{1, S_Q}^*(\{1, 2\}) = \frac{\sigma}{m+2\sigma} q_{1, S_Q}^*(\{1\})$ by (.22) and (.23). Also note that $q_{1, S_Q}^*(\{1\})$ can be derived by

letting $s = 1$ in (.11) to have

$$q_{1,SQ}^*({1}) = \frac{\sigma(a - \mu) + am + \sigma y_1}{2b(m + \sigma)} \quad (.24)$$

Multiplying both sides of (.24) by $\frac{\sigma}{m+2\sigma}$ gives $q_{1,SQ}^*(N)$ as claimed.

Case 2: Firm n 's best reply:

Let $s = n$ and $S_1 = N$ in the pre-entry market analysis. Best replies of firms n and $n - 1$ are respectively given from (.11) and (.14) as

$$q_{n,SQ}^*(N) = \frac{a - \mu - bQ_{N \setminus \{n\}}(N) + E(u|\mathbf{y}_n)}{2b} \quad (.25)$$

and

$$q_{n-1,SQ}^*(N) = \frac{\sigma(a - \mu - bQ_{N \setminus \{n-1,n\}}(N) + E(u|\mathbf{y}_{n-1}))}{2b(m + \sigma n)} \quad (.26)$$

Rearranging (.26) gives

$$E(u|\mathbf{y}_{n-1}) = \frac{2b(m + \sigma n)q_{n-1,SQ}^*(N)}{\sigma} - (a - bQ_{N \setminus \{n-1,n\}}(N) - \mu) \quad (.27)$$

Finally note that

$$E(u|\mathbf{y}_n) = \frac{\sigma y_n}{m + \sigma n} + \frac{m + \sigma(n - 1)}{m + \sigma n} E(u|\mathbf{y}_{n-1}) \quad (.28)$$

by Lemma 1(i). First substitute (.27) into (.28) and then put the resulting value of $E[u|\mathbf{y}_n]$ into (.25) to have

$$q_{n,SQ}^*(N) = \frac{b(m+n\sigma)(2m+(2n-3)\sigma)q_{n-1}(N) - b\sigma^2 \sum_{j \leq n-2} q_j(N) + \sigma^2(a-\mu) + \sigma^2 y_n}{2b\sigma(m+n\sigma)} \quad (.29)$$

as desired.

Case 3: Best replies of firms in $\{2, 3, \dots, n - 1\}$:

Consider any two non-empty ordered subsets of N , i.e., $S = \{1, 2, \dots, s\}$ and $S' = \{1, 2, \dots, s'\}$ with $s, s' \in \{3, 4, \dots, n\}$ and $s \neq s'$. W.L.O.G. let $s' \leq s$. For each $i \in S \cap S' \setminus \{s'\}$, $q_{i,SQ}^*(S) = q_{i,SQ}^*(S')$ according to the logic given in (.22).

Therefore, for each $j \in N \setminus \{1, n\}$, firm $j - 1$ and j 's best replies in the market of N firms can be produced from (.18) as

$$q_{j-1,SQ}^*({1, 2, \dots, j}) = q_{j-1,SQ}^*(N) = \frac{\sigma(a - \mu - bQ_{N \setminus \{j-1, j, \dots, n\}}(N) + E(u|\mathbf{y}_{j-1}))}{2b(m + \sigma j)} \quad (.30)$$

$$q_{j,SQ}^*({1, 2, \dots, j + 1}) = q_{j,SQ}^*(N) = \frac{\sigma(a - \mu - bQ_{N \setminus \{j, j+1, \dots, n\}}(N) + E(u|\mathbf{y}_j))}{2b(m + \sigma(j + 1))} \quad (.31)$$

Rearranging (.30) gives

$$E(u|\mathbf{y}_{j-1}) = \frac{2b(m + \sigma j)q_{j-1,SQ}^*(N)}{\sigma} - (a - \mu - bQ_{N \setminus \{j, j+1, \dots, n\}}(N)) \quad (.32)$$

But by Lemma 1(i),

$$E(u|\mathbf{y}_j) = \frac{\sigma y_j}{m + \sigma j} + \frac{m + \sigma(j - 1)}{m + \sigma j} E(u|\mathbf{y}_{j-1}) \quad (.33)$$

First put (.32) into (.33) and then replace the resulting value of $E(u|\mathbf{y}_j)$ in (.31) to have

$$q_{j,SQ}^*(N) = \frac{b(m + j\sigma)(2m + (2j - 3)\sigma)q_{j-1}(N) - b\sigma^2 \sum_{k \leq j-2} q_k(N) + \sigma^2(a - \mu) + \sigma^2 y_j}{2b(m + j\sigma)(m + (j + 1)\sigma)} \quad (.34)$$

as desired.

Step 4) Second Order Conditions

We finally prove that the second order conditions are satisfied, i.e., for all $i \in N$, $\frac{\partial^2 E[\pi_i | y_i, q_1, \dots, q_{i-1}]}{\partial^2 q_i} < 0$. Note first that as $b > 0$, $\frac{\partial^2 E[\pi_n | \cdot]}{\partial^2 q_n} = -2b < 0$ by (.9) for $s = n$. For $i < n$, the profit function of firm i is:

$$\begin{aligned} E[\pi_i(q_1, \dots, q_i, q_{i+1,SQ}^*, \dots, q_{n,SQ}^*, u) | y_i, q_1, \dots, q_{i-1}] &= \\ &= E[(a - b \sum_{j=1}^{j=i} q_j - b \sum_{k=i+1}^{k=n} q_{k,SQ}^* + u - \mu)q_i | y_i, q_1, \dots, q_{i-1}] \end{aligned} \quad (.35)$$

The second order condition is:

$$\frac{\partial^2 E[\pi_i | \cdot]}{\partial^2 q_i} = -2b(1 + \sum_{j=i+1}^n \frac{\partial q_{j,SQ}^*}{\partial q_i}) \quad (.36)$$

First let $i = n - 1$. Plugging in the value of $\frac{\partial q_{n,SQ}^*}{\partial q_{n-1}}$ from (.29) into (.36) yields

$$\frac{\partial^2 E[\pi_{n-1}|\cdot]}{\partial^2 q_{n-1}} = -2b\left(1 + \frac{2m + (2n - 3)\sigma}{2\sigma}\right) \quad (.37)$$

which is negative for $n \geq 2$. Lastly, let $i \leq n - 2$. After finding the partial derivatives in (.36) by using (.29) and (.34), we plug them into (.36) to have

$$\frac{\partial^2 E[\pi_i|\cdot]}{\partial^2 q_i} = -2b\left(1 + \frac{2m+(2i-1)\sigma}{2(m+(i+2)\sigma)} - \sum_{j=i+2}^{n-1} \frac{\sigma^2}{(m+j\sigma)(m+(j+1)\sigma)} - \frac{\sigma}{2(m+n\sigma)}\right) \quad (.38)$$

Note that the second term in the parenthesis is positive for $m, \sigma \in R_+$. Therefore, it is sufficient to show that the absolute value of last two subtraction terms in (.38) is smaller than 1. The absolute value of this term is maximized at $m = 0$ and given by

$$\frac{1}{2n} + \sum_{j=i+2}^{n-1} \frac{1}{j(j+1)} = \frac{1}{2n} + \sum_{j=i+2}^{n-1} \left(\frac{1}{j} - \frac{1}{j+1}\right) \quad (.39)$$

The first term in (.39) gets a maximum value of $1/4$ when $n = 2$. The second term in (.39) is decreasing in j and increasing in n . As $i \geq 1$, the second is maximized at $j = 3$ and $n \rightarrow \infty$ and gets a maximum value of $1/3$. Altogether, the maximum value of (.39) is smaller than $1/4 + 1/3 = 7/12$, which is smaller than 1. This observation completes the proof of Theorem 2. \square

We next prove three technical results. We use the first result in the proof of the second result; and use the second result in the derivation of the third result.⁵⁹ These results will later be used in the proof of Theorem 3.

Lemma 5. *Let $m, \sigma \in \mathbb{R}_+$.*

i) $E(Q_{SQ}^(N)) = a/b - E(q_{n,SQ}^*(N))$.*

ii) For each $i \in N \setminus \{n\}$,

$$E(q_{i,SQ}^*(N)) = \frac{a\sigma \prod_{j \in \{3,5,\dots,2i-1\}} (2m + j\sigma)}{b2^i \prod_{j \in \{2,3,4,\dots,i+1\}} (m + j\sigma)} \quad (.40)$$

⁵⁹As we use some derivations from the proof of Theorem 2 in the proofs of the following lemma, we did not want to put this lemma together with other lemmas in the Appendix.

For $i = n$,

$$E(q_{n,SQ}^*(N)) = \frac{a \prod_{j \in \{3,5,\dots,2n-1\}} (2m + j\sigma)}{b2^n \prod_{j \in \{2,3,4,\dots,n\}} (m + j\sigma)} \quad (.41)$$

iii) For any $n \geq 2$, $\frac{\partial E(q_{n,SQ}^*(N))}{\partial m} > 0$.

Proof of Lemma 5:

Proof of Lemma 5(i): We claim that $E(Q_{SQ}^*(N)) = a/b - E(q_{n,SQ}^*(N))$. Recall that $bQ_{SQ}^*(N) = a - bq_{n,SQ}^*(N) - \mu + E(u|\mathbf{y}_n)$ from the first order condition of the last follower given by (.10) for $s = n$. Taking expectations of both sides of this equality yields

$$bE(Q_{SQ}^*(N)) = a - bE(q_{n,SQ}^*(N)) - \mu + E(E(u|\mathbf{y}_n)) \quad (.42)$$

Since $E(E(u|\mathbf{y}_n)) = \mu$ by the law of iterated expectations, (.42) reduces to our claim as desired.

Proof of Lemma 5(ii): Let $n \geq 1$ and $m, \sigma \in \mathbb{R}_+$. Take any non-empty $S_1, S_2 \subset N$ with $S_2 \equiv S_1 \cup \{s+1\} \equiv \{1, 2, \dots, s, s+1\}$. Note that

$$E[q_{s,SQ}^*(S_2)] = \frac{\sigma}{m + \sigma(s+1)} E[q_{s,SQ}^*(S_1)] \quad (.43)$$

from (.23). Moreover, for each $S \in \{S_1, S_2\}$, we have

$$E[Q_{SQ}^*(S)] = a/b - E[q_{s,SQ}^*(S)] \quad (.44)$$

from Lemma 5(i). Finally, note that for each $i \in S_1 \setminus \{s\}$,

$$q_{i,SQ}^*(S_1) = q_{i,SQ}^*(S_2) \quad (.45)$$

from (.22). Thus, subtracting $E[Q_{SQ}^*(S_1)]$ from $E[Q_{SQ}^*(S_2)]$ by using (.44) and substituting the value of $E[q_{s,SQ}^*(S_2)]$ from (.43) into the resulting equation, rear-

ranging terms yields

$$E[q_{s+1,SQ}^*(S_2)] = \frac{(2m + \sigma(2s + 1))E[q_{s,SQ}^*(S_1)]}{2(m + \sigma(s + 1))} \quad (.46)$$

But remark that the expected monopoly output is given by

$$E[q_{1,SQ}^*({1})] = \frac{a}{2b} \quad (.47)$$

from (.24). Hence, starting from the monopoly market, when we allow entry of new firms as last-movers one by one until reaching n , then we can iteratively calculate the expected quantities of firms by using (.43), (.45), and (.46) to have (.40) and (.41). Hence all expected quantities are strictly positive for any $m, \sigma \in \mathbb{R}_+$.

Proof of Lemma 5(iii): We prove that $\frac{\partial E(q_{n,SQ}^*(N))}{\partial m} > 0$ by induction. Let $N_i = \{1, 2, \dots, i\}$. Note first that $\frac{\partial E(q_{2,SQ}^*(N_2))}{\partial m} = \frac{a\sigma}{4b(m+2\sigma)^2} > 0$ by (.41). Now assume that $\frac{\partial E(q_{k,SQ}^*(N_k))}{\partial m} > 0$ for some $k \in \{2, 3, \dots, n-1\}$. We want to show that $\frac{\partial E(q_{k+1,SQ}^*(N_{k+1}))}{\partial m} > 0$ as well. Derive from (.41) that $E(q_{k+1,SQ}^*(N_{k+1})) = \frac{2m+(2k+1)\sigma}{2m+2(k+1)\sigma} E(q_{k,SQ}^*(N_k))$. Derivating both sides of this equality with respect to m gives

$$\frac{\partial E(q_{k+1,SQ}^*(N_{k+1}))}{\partial m} = \frac{2m+(2k+1)\sigma}{2m+2(k+1)\sigma} \frac{\partial E(q_{k,SQ}^*(N_k))}{\partial m} + \frac{\sigma}{2(m+(k+1)\sigma)^2} E(q_{k,SQ}^*(N_k)) \quad (.48)$$

As $\frac{\partial E(q_{k,SQ}^*(N_k))}{\partial m} > 0$ by the initial assumption and $E(q_{k,SQ}^*(N_k)) > 0$ at $m, \sigma \in \mathbb{R}_+$ by (.41), then $\frac{\partial E(q_{k+1,SQ}^*(N_{k+1}))}{\partial m} > 0$ from (.48) as desired. \square

Proof of Theorem 3:

The expected monopoly output is the same under both competitions and is given by $E(Q_C^*({1})) = E(Q_{SQ}^*({1})) = a/2$ from (.40). Therefore, let $n \geq 2$. Remark that $E(Q_{SQ}^*(N)) = a/b - E(q_{n,SQ}^*(N))$ by Lemma 5(i). Together with (11), we have $\Delta E(Q(N)) = E(Q_C^*(N)) - E(Q_{SQ}^*(N)) = E(q_{n,SQ}^*(N)) - \frac{a}{b(n+1)}$. Since $\frac{\partial E(q_{n,SQ}^*(N))}{\partial m} > 0$ by Lemma 5(iii), it is sufficient to consider $m = 0$. Plugging in the value of $E(q_{n,SQ}^*(N))$ from (.41) into $\Delta E(Q(N))$ and evaluating the resulting

equation at $m = 0$ gives

$$\Delta E(Q(N)) = \frac{a}{b} \left(\frac{(2n-1)!!}{(2n)!!} - \frac{1}{n+1} \right) \quad (.49)$$

where $(2n-1)!! = 1 * 3 * \dots * (2n-3) * (2n-1)$ and $(2n)!! = 2 * 4 * \dots * (2n-2) * (2n)$. We prove that (.49) is positive by induction. Let $N_i = \{1, 2, \dots, i\}$. Note that $\Delta E(Q(N_2)) = a/(24b) > 0$ by (.49). Now suppose that $\Delta E(Q(N_k)) > 0$ for some $k \in \{2, 3, \dots, n-1\}$. We want to show that $\Delta E(Q(N_{k+1})) > 0$ as well. Simple algebra shows that $\frac{1}{k+1} - \frac{2(k+1)}{(k+2)(2k+1)} = \frac{k}{(k+1)(k+2)(2k+1)} > 0$. Since $\frac{(2k-1)!!}{(2k)!!} > \frac{1}{k+1}$ by the initial supposition that $\Delta E(Q(N_k)) > 0$, we then have $\frac{(2k-1)!!}{(2k)!!} > \frac{2(k+1)}{(k+2)(2k+1)}$ as well. But that is equivalent to say that $E(Q(N_{k+1})) > 0$ by (.49) as desired. Hence the claim is proven. □

Proof of Theorem 4:

Let $a, m, \sigma \in \mathbb{R}_+$. Define $\Delta E(\Pi^*(N)) = nE(\pi_C^*(N)) - \sum_{i \in N} E(\pi_{i,SQ}^*(N))$. The leader and follower's profits in a Stackelberg duopoly are respectively given by

$$E(\pi_{1,SQ}^*(\{1, 2\})) = \frac{\sigma(2m+3\sigma)(a^2(m+\sigma)+\sigma^2)}{8b(m+\sigma)(m+2\sigma)^2} \quad (.50)$$

$$E(\pi_{2,SQ}^*(\{1, 2\})) = \frac{a^2(m+\sigma)(2m+3\sigma)^2 + \sigma^2(8m^2+20m\sigma+9\sigma^2)}{16b(m+\sigma)(m+2\sigma)^2} \quad (.51)$$

based on (4), (16), and equilibrium quantities of the leader and follower provided by Theorem 2. Similarly, using Theorem 1, expected per firm Cournot profits are symmetric:

$$E[\pi_{i,C}^*(N)] = E[\pi_C^*(N)] = \frac{a^2}{b(n+1)^2} + \frac{\sigma^2(m+\sigma)}{b(2m+\sigma(n+1))^2} \quad (.52)$$

First sum up (.50) and (.51) to have total Stackelberg duopoly equilibrium expected profits and then subtract the summation from $2E[\pi_C^*(N)]$ to get

$$\Delta E(\Pi^*({1, 2})) = -\frac{a^2(m+\sigma)(2m+\sigma)(2m+3\sigma)^2(2m+7\sigma)+9\sigma^4(4m^2+12m\sigma+7\sigma^2)}{144b(m+\sigma)(m+2\sigma)^2(2m+3\sigma)^2} < 0$$

Similar calculations at $n = 3, 4, 5$ yield

$$\begin{aligned} \Delta E(\Pi^*({1, 2, 3})) &= -\frac{a^2(m+\sigma)(2m^2+6m\sigma+3\sigma^2)(2m^2+14m\sigma+21\sigma^2)+7\sigma^4(4m^2+16m\sigma+9\sigma^2)}{64b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2} < 0 \\ \Delta E(\Pi^*({1, 2, 3, 4})) &= -\frac{9a^2(m+\sigma)(2m+5\sigma)^2(8m^3+52m^2\sigma+98m\sigma^2+47\sigma^3)(8m^3+92m^2\sigma+318m\sigma^2+337\sigma^3)}{6400b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(2m+5\sigma)^2} \\ &\quad -\frac{25\sigma^4(1408m^6+21184m^5\sigma+127072m^4\sigma^2+386752m^3\sigma^3+623020m^2\sigma^4+493000m\sigma^5+142551\sigma^6)}{6400b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(2m+5\sigma)^2} < 0 \\ \Delta E(\Pi^*({1, 2, 3, 4, 5})) &= -\frac{a^2(m+\sigma)(32m^4+352m^3\sigma+1336m^2\sigma^2+2000m\sigma^3+915\sigma^4)(32m^4+544m^3\sigma)}{9216b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(m+5\sigma)^2} \\ &\quad -\frac{a^2(m+\sigma)(32m^4+352m^3\sigma+1336m^2\sigma^2+2000m\sigma^3+915\sigma^4)(3208m^2\sigma^2+7856m\sigma^3+6765\sigma^4)}{9216b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(m+5\sigma)^2} \\ &\quad -\frac{9\sigma^4(3200m^6+57920m^5\sigma+411136m^4\sigma^2+1448560m^3\sigma^3+2628668m^2\sigma^4+2271864m\sigma^5+687775\sigma^6)}{9216b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(m+5\sigma)^2} < 0 \end{aligned}$$

□

Proof of Theorem 5:

Let $a, m, \sigma \in \mathbb{R}_+$. Let $\Delta E(TW^*(N)) = E(TW_C^*(N)) - E(TW_{SQ}^*(N))$ and $\Delta E(CS^*(N)) = E(CS_C^*(N)) - E(CS_{SQ}^*(N))$. Based on (4), (17), and equilibrium quantities of Cournot and Stackelberg games, which are respectively given by Theorems 1 and 2, total duopoly expected equilibrium welfare under both competitions are derived as

$$E(TW_C^*({1, 2})) = \frac{4a^2(2m+3\sigma)^2+9\sigma^2(3m+4\sigma)}{9b(2m+3\sigma)^2} \quad (.53)$$

$$E(TW_S^*({1, 2})) = \frac{a^2(m+\sigma)(2m+5\sigma)(6m+11\sigma)+\sigma^2(24m^2+76m\sigma+55\sigma^2)}{32b(m+\sigma)(m+2\sigma)^2} \quad (.54)$$

Subtracting (.54) from (.53) yields

$$\Delta E(TW^*({1, 2})) = \frac{a^2(m+\sigma)(2m+\sigma)(2m+3\sigma)^2(10m+17\sigma)+9\sigma^3(16m^3+60m^2\sigma+64m\sigma^2+17\sigma^3)}{288b(m+\sigma)(m+2\sigma)^2(2m+3\sigma)^2} > 0$$

Similar calculations at $n = 3, 4, 5$ give

$$\begin{aligned}
\Delta E(TW^*({1,2,3})) &= \frac{a^2(m+\sigma)(2m^2+6m\sigma+3\sigma^2)(6m^2+26m\sigma+27\sigma^2)+s^3(48m^3+236m^2\sigma+296m\sigma^2+81\sigma^3)}{128b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2} > 0 \\
\Delta E(TW^*({1,2,3,4})) &= \frac{25\sigma^2(512m^8+11776m^7\sigma+114816m^6\sigma^2+620352m^5\sigma^3)}{6400b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(2m+5\sigma)^2} + \\
&+ \frac{25\sigma^2(2033568m^4\sigma^4+4145472m^3\sigma^5+5135956m^2\sigma^6+3537464m\sigma^7+1037097\sigma^8)}{6400b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(2m+5\sigma)^2} \\
&+ \frac{a^2(m+\sigma)(2m+5\sigma)^2(1472m^6+26496m^5\sigma+199376m^4\sigma^2+799152m^3\sigma^3+1792076m^2\sigma^4+2124156m\sigma^5+1037097\sigma^6)}{6400b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(2m+5\sigma)^2} > 0 \\
\Delta E(TW^*({1,2,3,4,5})) &= \frac{9\sigma(89039120m^3\sigma^6+65894692m^2\sigma^7+20927176m\sigma^8+483425\sigma^9)}{18432b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(m+5\sigma)^2} + \\
&+ \frac{9\sigma(5120m^9+129280m^8\sigma+1388800m^7\sigma^2+8270720m^6\sigma^3+29794560m^5\sigma^4+66343584m^4\sigma^5)}{18432b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(m+5\sigma)^2} + \\
&+ \frac{a^2(m+\sigma)(32m^4+352m^3\sigma+1336m^2\sigma^2+2000m\sigma^3+915\sigma^4)(64m^4+800m^3\sigma+3608m^2\sigma^2+6928m\sigma^3+4755\sigma^4)}{18432b(m+\sigma)(m+2\sigma)^2(m+3\sigma)^2(m+4\sigma)^2(m+5\sigma)^2} > 0
\end{aligned}$$

But since $\Delta E(TW^*(N)) = \Delta E(CS^*(N)) + \Delta E(\Pi^*(N))$, then it becomes a corollary that $\Delta E(CS^*(N)) > 0$ at $n = 2, 3, 4, 5$ by Theorem 4 and the above total welfare analysis. \square

Proof of Theorem 6: Let $n = 2$ and $c_1 = c_2 = c$ be the pre-merger marginal cost level. Let also $a = \bar{a} - c$. Let $\bar{c}_{e,C}$ and $\bar{c}_{e,S}$ be the minimum level of efficiency gains under Cournot and Stackelberg competitions respectively such that the merger to the monopoly is total welfare enhancing. Accordingly, let $\hat{a}_S = \bar{a} - c + \bar{c}_{e,S}$ and $\hat{a}_C = \bar{a} - c + \bar{c}_{e,C}$. We claim that $\bar{c}_{e,C} > \bar{c}_{e,S}$. To see that note that \hat{a}_C solves $TW_C^*\{1, 2\} = TW_M^*$ and therefore is given by

$$\hat{a}_C = a\sqrt{\frac{32}{27b^2}} + \sqrt{\frac{18\sigma^3(4m+5\sigma)}{27b^2(m+2\sigma)(2m+3\sigma)^2}} \quad (.55)$$

where TW_M^* and $TW_C^*\{1, 2\}$ are respectively provided in (29) and (.53). Similarly, \hat{a}_S solves $TW_S^*\{1, 2\} = TW_M^*$ and is given by

$$\hat{a}_S = a\sqrt{\frac{(2m+5\sigma)(6m+11\sigma)}{12b^2(m+2\sigma)^2}} + \sqrt{\frac{\sigma^3(4m+7\sigma)}{12b^2(m+\sigma)(m+2\sigma)^2}} \quad (.56)$$

where $TW_S^*\{1, 2\}$ is given by (.54). Subtracting (.56) from (.55), after some simplifications, reduces to

$$a\sqrt{\frac{(2m+\sigma)(10m+17\sigma)}{108b^2(m+2\sigma)^2}} + \sqrt{\frac{\sigma^3(16m^3+60m^2\sigma+64m\sigma^2+17\sigma^3)}{12b^2(m+\sigma)(m+2\sigma)^2(2m+3\sigma)^2}} > 0 \quad (.57)$$

which is positive for $m, \sigma \in \mathbb{R}_+$. In sum, $\hat{a}_C > \hat{a}_S$. This implies that $\bar{c}_{e,C} > \bar{c}_{e,S}$ as desired. \square

Proof of Theorem 8: The proof is very similar to the proof of Proposition 2 in Vives (1984). We first show that if firm one uses $q_{1,DC}^*(\{1, 2\}) = \frac{a}{b(2+\lambda)} + \frac{\sigma(y_1 - \mu)}{b(2m + \sigma(2+\lambda))}$, then the unique best response for firm 2 is to use $q_{2,DC}^*(\{1, 2\}) = \frac{a}{b(2+\lambda)} + \frac{\sigma(y_2 - \mu)}{b(2m + \sigma(2+\lambda))}$. To see this, notice that the expected profit of firm 2 choosing quantity strategy q_2 given firm one's strategy and the signal y_2 is

$$E[(a - \mu + u - bq_2 - bq_{1,DC}^*(\{1, 2\}))q_2|y_2] \quad (.58)$$

Therefore, the optimal choice of firm 2 is

$$q_{2,DC}^*(\{1, 2\}) = \frac{a - \mu + E(u|y_2) - bE(q_{1,DC}^*(N)|y_2)}{2b} \quad (.59)$$

which, after some computations, is equal to $q_{2,DC}^*(\{1, 2\}) = \frac{a}{b(2+\lambda)} + \frac{\sigma(y_2 - \mu)}{b(2m + \sigma(2+\lambda))}$ by the initial supposition and Lemma 1 (at $n = 2$). Uniqueness follows similarly as in Vives (1984). \square

Proof of Theorem 8: We proceed in two steps to derive the equilibrium quantities of firms, who sell differentiated products.

Step 1: Deriving Equilibrium Productions:

We start out with writing down firm two's maximization problem:

$$\max_{q_2} E[\pi_2(q_1, q_2, u)|y_2, q_1]$$

The first-order condition yields

$$q_{2,DS}^* = \frac{a - \mu - b\lambda q_1 + E(u|y_2, q_1)}{2b} \quad (.60)$$

The inverse functions of firm strategies $G_1(y_1)$ and $G_2(y_2, q_1)$ exist by the definition of equilibrium and are linear by our conjecture. In essence, firm two can perfectly infer y_1 after observing q_1 from the definition of perfect revealing equilibrium. Hence substituting $E[u|y_2, q_1] = \frac{\sigma(y_1 + y_2)}{m + 2\sigma} + \frac{m\mu}{m + 2\sigma}$, where $y_1 = G_1^{-1}(q_1) = \frac{q_1 - \alpha_0}{\alpha_1}$, in (.60) by Lemma 1(i) and rearranging terms give

$$q_{2,DS}^* = G_1(y_2, q_1) = \frac{a}{2b} - \frac{\sigma(\alpha_0 + 2\alpha_1\mu)}{2b\alpha_1(m + 2\sigma)} + \frac{\sigma y_2}{2b(m + 2\sigma)} + \left(\frac{-\lambda}{2} + \frac{\sigma}{2b\alpha_1(m + 2\sigma)}\right)q_1$$

We first substitute $q_{2,DS}^*$ from above in the leader firm's optimization problem as

$$\max_{q_1} E[\pi_1(q_1, q_{2,DS}^*, u)|y_1] = E[(a - \mu + u - bq_1 - b\lambda q_{2,DS}^*)q_1|y_1] \quad (.61)$$

After further substituting $E[u|y_1] = E[y_2|y_1] = \frac{\sigma y_1}{m+\sigma} + \frac{\sigma\mu}{m+\sigma}$ in (.61) by Lemma 1, one can solve for α_0 and α_1 from the first order conditions for the maximization problem in (.61) and obtain equilibrium quantities of both firms.

STEP 2: Second Order Conditions for the Differentiated Good Model

Observe first that $\frac{\partial^2 E[\pi_2(\cdot)|y_2, q_1]}{\partial^2 q_2} = -2b < 0$ and the second order condition for the follower is satisfied.

Finally note that $\frac{\partial^2 E[\pi_1(\cdot)|y_1]}{\partial^2 q_1} = -2b(1 + \lambda\beta_3)$ by (.61). As $\beta_3 = \frac{\partial^2 q_2^*}{\partial^2 q_1} = \frac{(2-\lambda)(2m+(2-\lambda)\sigma)}{4(1-\lambda)m+2(4-3\lambda)\sigma} > 0$ for any $\lambda \in [0, 1]$ and $m, \sigma \in \mathbb{R}_+$ from Theorem 8, $\frac{\partial^2 E[\pi_1(\cdot)|y_1]}{\partial^2 q_1} < 0$. Hence, all second order conditions are satisfied. \square

Proof of Theorem 9:

i) Let $E[\Delta Q_D^*] = E[Q_{DC}^*] - E[Q_{DS}^*]$. By using (4) and equilibrium quantities of firms in differentiated Cournot and Stackelberg models respectively from Theorems (7) and (8), it can be shown that

$$E[\Delta Q_D^*] = \frac{a\lambda(2-\lambda)(2m + \sigma(2-\lambda))}{4b(2+\lambda)(2-\lambda^2)(m+2\sigma)} \quad (.62)$$

which is positive for $\lambda \in (0, 1]$ and is equal to zero when $\lambda = 0$.

ii) Let $a \geq \sigma$ and $\lambda \in [0, 1]$. We claim that $E[\Delta \Pi_D^*] = E[\Pi_{DC}^*] - E[\Pi_{DS}^*] < 0$. By using (4) and the equilibrium quantities of firms in differentiated Cournot and Stackelberg models respectively stated in Theorems (7) and (8), it can be shown that

$$\frac{\partial E[\Delta \Pi_D^*]}{\partial a} = -\frac{2a(m(2(1-\lambda)(4-\lambda^2) + 2\lambda^3) + \sigma(2-\lambda)(12-5\lambda^2))}{g} < 0 \quad (.63)$$

at $\lambda \in [0, 1]$ and $m, \sigma \in \mathbb{R}_+$ where $g = 16b(2 + \lambda)^2(2 - \lambda^2)^2(m + \sigma)(m + 2\sigma)^2(2m + 2\sigma + \lambda\sigma)^2$. Therefore, it is sufficient to prove that $E[\Delta\Pi_D^*(a = \sqrt{\sigma})] < 0$. Calculating $E[\Delta\Pi_D^*(.)]$ at $a = \sqrt{\sigma}$ and rearranging terms yields

$$\begin{aligned} E[\Delta\Pi_D^*(a = \sqrt{\sigma})] &= -\frac{\sigma(m+2\sigma)(16\lambda^2(\lambda^3+(1-\lambda)(4-\lambda^2))m^4+32(8(1-\lambda)+2\lambda^5+\lambda^2(2-\lambda^2))m^3\sigma)}{\dots} \\ &\quad -\frac{\sigma(m+2\sigma)(8(16\lambda^2+32(1-\lambda)(2-\lambda-\lambda^2)+\lambda^3(2+(3+\lambda)(10(1-\lambda)+\lambda^3)))m^2\sigma^2)}{\dots} \\ &\quad -\frac{\sigma(m+2\sigma)(8(2-\lambda)(2+\lambda)(8(1-\lambda)^2+2\lambda^2+(6-5\lambda)\lambda^3+2\lambda^3(3-\lambda^2))m\sigma^3+\lambda^2(12-5\lambda^2)(4-\lambda^2)^2\sigma^4)}{g} \end{aligned}$$

which is negative at $\lambda \in [0, 1]$ and $m, \sigma \in \mathbb{R}_+$ as desired. \square

Proof of Theorem 10: Let $m, \sigma \in \mathbb{R}_+$. Let $E[\Delta TW_D^*] = E[TW_{DC}^*] - E[TW_{DS}^*]$. By using (4) and the equilibrium quantities of firms in differentiated Cournot and Stackelberg models respectively stated in Theorems (7) and (8), it can be shown that the partial derivative $\frac{\partial \Delta E[TW_D^*]}{\partial a}$ reduces to

$$\frac{2a(8(2 - \lambda)m + 4(1 - \lambda^2)(9\sigma + 4m) + 4\sigma + 24\sigma(1 - \lambda) + 3\sigma\lambda^4 + \lambda^3(2m + 10\sigma))}{32b(2 + \lambda)^2(2 - \lambda^2)^2(m + \sigma)(m + 2\sigma)^2(2m + \sigma(2 + \lambda))^2}$$

which is positive at $m, \sigma \in \mathbb{R}_+$. In that regard, it is sufficient to show that $E[\Delta TW_D^*(a = 3\sqrt{\sigma})] > 0$ at $\lambda \in [0.1, 1]$ to conclude the proof. Substituting $a = 3\sqrt{\sigma}$ into $E[\Delta TW_D^*]$ and rearranging terms yield

$$\begin{aligned} E[\Delta TW_D^*(a=3\sqrt{\sigma})] &= \frac{144\lambda(4+4(1-\lambda)+\lambda^3+8(1-\lambda^2))m^5\sigma+16(f_1+\lambda^4(22+\lambda(22+3\lambda)))m^4\sigma^2}{32b(2+\lambda)^2(2-\lambda^2)^2(m+\sigma)(m+2\sigma)^2(2m+2\sigma+\lambda\sigma)^2} + \\ &\quad + \frac{8(f_2+\lambda(76(1-\lambda^3)+\lambda^4(268+\lambda(33-2\lambda))))m^3\sigma^3+4(f_3+\lambda 604(1-\lambda^3)+\lambda^5(1204+\lambda(1+34(1-\lambda)+3(35-\lambda^2))))m^2\sigma^4}{32b(2+\lambda)^2(2-\lambda^2)^2(m+\sigma)(m+2\sigma)^2(2m+2\sigma+\lambda\sigma)^2} + \\ &\quad + \frac{(2-\lambda)(2+\lambda)(f_4+496\lambda(1-\lambda^3)+3\lambda^5(96+17\lambda))m\sigma^5+10\lambda(4-\lambda^2)^2(16(2-\lambda^2)+\lambda(4-3\lambda^2))\sigma^6}{32b(2+\lambda)^2(2-\lambda^2)^2(m+\sigma)(m+2\sigma)^2(2m+2\sigma+\lambda\sigma)^2} \end{aligned} \tag{.64}$$

where

$$\begin{aligned} f_1 &= -48 + 896\lambda - 124\lambda^2 - 492\lambda^3 \\ f_2 &= -384 + 4276\lambda - 128\lambda^2 - 2704\lambda^3 \\ f_3 &= -960 + 9700\lambda + 608\lambda^2 - 7424\lambda^3 \\ f_4 &= -384 + 5392\lambda + 624\lambda^2 - 3520\lambda^3 \end{aligned}$$

Note that $\frac{\partial^2 f_1}{\partial^2 \lambda} = -8(31 + 369\lambda)$, $\frac{\partial^2 f_2}{\partial^2 \lambda} = -32(8 + 507\lambda)$, $\frac{\partial^2 f_3}{\partial^2 \lambda} = 64(19 - 696\lambda)$, and $\frac{\partial^2 f_4}{\partial^2 \lambda} = 96(13 - 220\lambda)$. Therefore, all f functions are concave at $\lambda \geq 0.1$. Furthermore, at $\lambda = 0.1$, $f_1 = 39.9$, $f_2 = 39.6$, $f_3 = 8.66$, and $f_4 = 157.9$, which are all positive. Similarly, at $\lambda = 1$, $f_1 = 232$, $f_2 = 1060$, $f_3 = 1924$, and

$f_4 = 2112$, which are also positive. Altogether, all f functions are positive when $\lambda \in [0.1, 1]$ by concavity. It is then easy to see from (.64) that $E[\Delta TW_D^*(a = 3\sqrt{\sigma})] > 0$ at $\lambda \in [0.1, 1]$.

By using Theorem 9(ii) and the above results, it is trivial to see that $E[\Delta CS_D^*] > 0$ at $a \geq 3\sigma$ and $\frac{\partial \Delta E[CS_D^*]}{\partial a} > 0$.

□

References

- [1] ABELSON, R. (2004), “Makers of Drug-coated stents fight for the hearts of cardiologists”, *New York Times*, May 12.
- [2] ALBAEK, S. (1990), “Stackelberg Leadership as a Natural Solution Under Cost Uncertainty”, *The Journal of Industrial Economics*, **38**, No. 3., 335-347.
- [3] ANDERSON, S. P. and M. ENGERS (1992), “Stackelberg versus Cournot Oligopoly Equilibrium”, *International Journal of Industrial Organization*, **10**, 27-135.
- [4] BAGWELL, K. (1995), “Commitment and Observability in Games”, *Games and Economic Behaviour*, **8**, 271-280.
- [5] BOYER, M. and M. MOREAUX (1986), “Perfect Competition as the Limit of a Hierarchical Market Game”, *Economic Letters*, **22**, 115-118. North Holland.
- [6] BOYER, M. and M. MOREAUX (1987), “On Stackelberg Equilibria with Differentiated Products: The Critical Role of the Strategy Space”, *Journal of Industrial Economics*, **36**, 217-230.
- [7] CUMBUL, E. (2014), “*First-mover and Last-mover Advantages with Private Information*”, Working Paper, IESE Business School.
- [8] DAUGHETY, F. A. (1990), “Beneficial Concentration”, *American Economic Review*, **80**, No. 5., 1231-1237.
- [9] DEGROOT, M. H. (1970), “*Optimal Statistical Decisions*”, Chapter 9, McGraw-Hill, New York.
- [10] DEMIGUEL, V. and H. XU (2009), “A Stochastic Multiple-Leader Stackelberg Model: Analysis, Computation, and Application”, “*Operations Research*”, **57**, 5, 1220-1235.
- [11] DE WOLF, D. and Y. SMEERS (1997), “A Stochastic Version of a Stackelberg-Nash-Cournot Equilibrium Model”, “*Management Science*”, **43**, 2, 190-197.
- [12] DISTASO W., P. LUPI, and F. M. MANENTI (2006), “Platform Competition and Broadband Uptake: Theory and Empirical Evidence from the European Union”, “*Information Economics and Policy*”, **18**, 2, 87-106.
- [13] DrameXchange, (May 2012), “*Trendforce: Global Branded Mobile DRAM Manufacturer 1Q12 Revenue Ranking-Samsung Takes Nearly 60% of the Market*”, www.dramexchange.com/WeeklyResearch/Post/5/3034.html

- [14] ERICSON, W. A. (1969), “A Note on the Posterior Mean of a Population Mean”, *Journal of the Royal Statistical Society*, **31**, 332-334.
- [15] FLAMM, K. (2005), “The Role of Economics, Demographics, and State Policy in Broadband Availability”, LBJ School of Public Affairs: Austin, TX.
- [16] GAL-OR, E. (1987), “First-mover Disadvantages with Private Information”, *Review of Economic Studies*, **54**, 279-292.
- [17] GAL-OR, E., T. GEYLANI, and J. A. DUKES (2008), “Information Sharing in a Channel with Partially Informed Retailers”, *Marketing Science*, **27**, No 4, 642-658.
- [18] GAL-OR, E., T. GEYLANI, and J. A. DUKES (2011), “Who Benefits from Bilateral Information Exchange in a Retail Channel”, *Economics Letters*, **112**, 210-212.
- [19] GENTZOGLANIS A. and E. ARAVANTINOS (2008), “Forecast Models of Broadband Diffusion and Other Information Technologies”, *Communications & Strategies*, 73.
- [20] GHEMAWAT P., B. NALEBUFF (1985), “Exit”, *Rand Journal of Economics*, **16**(2), 184-194.
- [21] GÜTH, W., G. KIRCHSTEIGER, and K. RITZBERGER (1998), “Imperfectly Observable Commitments in n -Player Games,” *Games and Economic Behaviour*, **23**, 54-74.
- [22] HACKBARTH, D. and MIAO J. (2012), “The Dynamics of Mergers and Acquisitions in Oligopolistic Industries” *Journal of Economic Dynamics & Control*, **36**, 585-609.
- [23] INO, H. and T. MATSUMURA (2012), “How Many Firms Should Be Leaders? Beneficial Concentration Revisited,” *International Economic Review*, **53**, 4, 1323-1340.
- [24] KADIYALI, V., K. SUDHIR, and R. V. RAO (2001), “Structural Analysis of Competitive Behavior: New Empirical Industrial Organization Methods in Marketing,” *International Journal of Research in Marketing*, **18**, 161-186.
- [25] KANG, J. (2010), “A Study of the DRAM Industry,” MSc. in Management Studies Thesis at the MIT Sloan School of Management.
- [26] KÜHN K. and X. VIVES (1995), “*Information Exchanges Among Firms and Their Impact on Competition*”, Luxembourg: Office for Official Publications of the European Communities, ISBN 92-826-9705-3.

- [27] JANSSEN, C. W. M. and E. MAASLAND (1997), “*On the Unique D1 Equilibrium in the Stackelberg Model with Asymmetric Information*,” CentER for Economic Research, Tilburg University, Working Paper.
- [28] LI, L. (1985), “Cournot Oligopoly with Information Sharing”, *Rand Journal of Economics*, **16**, 4, 521-536.
- [29] LIU, Z. (2005), “Stackelberg Leadership with Demand Uncertainty”, *Managerial and Decision Economics*, **26**, 345-350.
- [30] LOPEZ, A. R., A. M. AZZAM, and C. LIRÓN-ESPAÑA (2002), “Market Power and/or Efficiency: A Structural Approach”, *Review of Industrial Organization*, **20**, 115-126.
- [31] MAGGI, G. (1999), “The Value of Commitment with Imperfect Observability and Private Information”, *Rand Journal of Economics*, **30**, 4, 555-574.
- [32] MAILATH, G. (1993), “Endogeneous Sequencing of Firm Decisions”, *Journal of Economic Theory*, **59**, 169-182.
- [33] MAILATH, G. (1989), “Endogeneous Sequencing of Firm Decisions”, mimeo. University of Pennsylvania, Philadelphia.
- [34] MILGROM, P. and J. ROBERTS (1982), “Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis”, *Econometrica*, **50**, 2, 443-459.
- [35] MORGAN, J. and F. VARDY (2007), “The Value of Commitment in Contests and Tournaments When Observation is Costly”, *Games and Economic Behaviour*, **60**, 326-338.
- [36] NEWMARK M. C. (2004), “Price-Concentration Studies: There You Go Again”, Prepared for the DOJ/FTC Merger Workshop, “Concentration and Market Shares” Panel
- [37] OKUGUCHI, K. (1999), “Cournot and Stackelberg Duopolies Revisited”, *Japanese of Economic Review*, **50**, No. 3.
- [38] RAJU, S. J. and A. ROY (2000), “Market Information and Firm Performance”, *Management Science*, **46**, 8, 1075-1084.
- [39] ROBSON, J. A. (1990), “Stackelberg and Marshall”, *American Economic Review*, **80**, 1, 1990.
- [40] SHINKAI, T. (2000), “Second Mover Disadvantages in a Three-Player Stackelberg Game with Private Information”, *Journal of Economic Theory*, **90**, 293-304.

- [41] SPENCE, M. (1973), “Job Market Signaling”, *Quarterly Journal of Economics*, **83**, 3, 355-374.
- [42] TIROLE, J. (1995), “*The Theory of Industrial Organization*”, Eighth printing, Chapter 11, 450-453.
- [43] VARDY, F. (2004), “The Value of Commitment in Stackelberg Games with Observation Costs”, *Games and Economic Behaviour*, **49**, 374-400. b
- [44] VAN DAMME, E. and S. HURKENS (1997), “Games with Imperfectly Observable Commitment”, *Games and Economic Behaviour*, **21**, 282-308.
- [45] VIVES, X. (1984), “Duopoly Information Equilibrium”, *Journal of Economic Theory*, **34**, 71-94.
- [46] VIVES, X. (1988), “Sequential Entry, Industry Structure and Welfare”, *European Economic Review*, **32**, 8, 1671-1687.
- [47] VIVES, X. (1999), “*Oligopoly Pricing: Old Ideas and New Tools*”, MIT Press Books, The MIT Press, edition 1, vol. 1.
- [48] VIVES, X. (2011), “Strategic Supply Function Competition With Private Information”, *Econometrica*, **79**, No. 6., 1919-1966.
- [49] WALLSTEN, S. (2007), “Whence Competition in Network Industries? Broadband and Unbundling Regulations in OECD Countries”, working paper, Technology Policy Institute.
- [50] XU, H. (2005), “An MPCC Approach for Stochastic Stackelberg-Nash-Cournot Equilibrium”, *Optimization*, **54**, No. 1, 27-57