Competitive Pricing in the Presence of Manipulable Information in Online Platforms

Harish Guda* Yuqi Yang† Hongmin Li‡

January 10, 2024
(Click here for the latest version)

Abstract

Problem Definition: To entice customers to purchase, sellers on online platforms often misrepresent the quality of their goods/services, e.g., by manipulating consumer opinion. We analyze an oligopoly where sellers, heterogeneous in their true quality, compete by jointly choosing their prices and the extent of manipulation.

Methodology: Non-Cooperative Game Theory, Choice Models, and Optimization.

Results: We solve for the unique equilibrium when price-setting firms can manipulate their perceived quality and characterize the set of sellers that manipulate in equilibrium. We identify an index called the *propensity to manipulate*, based on model primitives to identify the set of sellers who have greater incentive to manipulate, and show that the set of sellers that manipulate in equilibrium is upward-closed in the propensity to manipulate. The extant literature has been mixed in its findings on which sellers have greater incentive to manipulate. Our work helps reconcile the differing viewpoints in the extant literature by providing a unified perspective.

Managerial Implications: We demonstrate the practical relevance of our model by mapping it to an environment consisting of sellers who are differentiated in a star-rating system based on their true rating and the volume of ratings. Depending on a seller's rating and volume of ratings, we identify three distinct regions that arise: a cost-prohibitive region, a cost-dominant region, and a benefit-dominant region. The ability to map a seller to one of these regions allows platform managers to understand a seller's tendency to manipulate consumer opinion dynamically over time.

*W.P. Carey School of Business, Arizona State University. hguda@asu.edu
†W.P. Carey School of Business, Arizona State University. yyang450@asu.edu
‡W.P. Carey School of Business, Arizona State University. hongmin.li@asu.edu

. . . I went to buy a pair of wireless earbuds. After I purchased them I got an email . . . telling me that they would give me a free wireless charger if (and only if) I gave a 5 star review. I contacted Amazon about it and they said it was against their policy to do that but they were not going to investigate the matter.

— Customer reports on sellers' efforts to manipulate ratings (Crockett, 2019).

The seller is obviously incentivizing people to leave positive reviews. Does Amazon even care? I'm pretty sure nothing will happen to him and he'll keep outranking me because I guess I'm dumb enough to play by the rules.

— Seller complains on Amazon Seller Forums (Amazon Seller Central, 2019).

1 Introduction

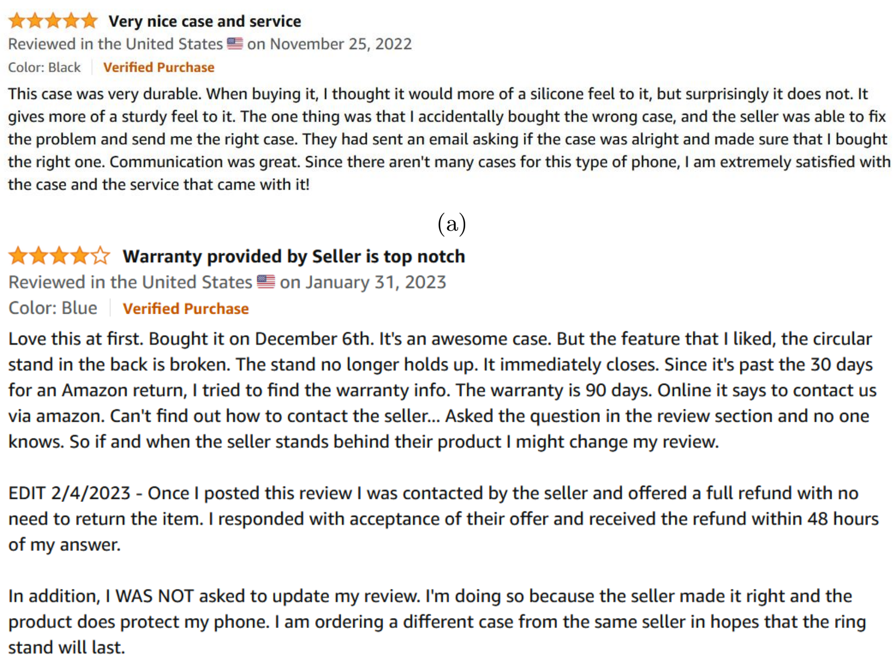
Internet-enabled marketplaces, e.g., retail platforms like Amazon and Ebay, provide consumers with the ability to not only engage in trade with sellers, but also provide a vast amount of information to help guide their purchasing decisions. Information on sellers' performance is typically user-generated in the form of consumer opinion or feedback, consisting of reviews and ratings, either on the platform or other product review forums. A vast literature, both in Marketing and Economics, has shown that consumers are influenced by such information in their purchase decisions (Chevalier and Mayzlin, 2006; Chintagunta et al., 2010; Mayzlin et al., 2014). Recent estimates by World Economic Forum (2021) suggest that consumer opinions via online reviews influence \$3.8 trillion of global commerce. In the context of restaurants, Luca (2011) estimates that a 1-star increase on Yelp rating leads to a 5-9% increase in revenues. Besides affecting consumers' purchase decisions, information on sellers' performance plays a critical role in the platform's listing strategy, e.g., in their search rankings. For instance, Amazon ranks sellers on various performance metrics, and awards the "buy-box" to their best performing sellers (Chen and Wilson, 2017).[1] This virtual word-of-mouth effect can form a reinforcing feedback loop that sets the sellers apart: those that succeed and those that fail.

Due to the competitive advantage that superior consumer opinion bestows on sellers, it is no surprise that sellers resort to manipulating these opinions via unfair means. A leading example through which sellers affect consumer opinion is *fake post-for-pay reviews*. In its simplest form, sellers solicit positive opinions that promote their products in exchange for a monetary transfer. While such manipulation of consumer opinion is often illegal, in a recent

[1] The buy-box refers to the white box on the right side of the Amazon product detail page, where customers add items for purchase to their cart. If left unchanged, Amazon assigns the default seller of a product to a top performing seller. Shoppers rarely browse a product's other sellers. Being awarded the buy-box is arguably one of the biggest perks a seller can get on the Amazon marketplace. It is estimated that 82% of a product's sales go through the buy-box. See https://www.bigcommerce.com/blog/win-amazon-buy-box/.

paper, He et al. (2022) show the existence of a large and active market for fake reviews. Recent estimates by certain large platforms show that 4% of online reviews are fake (World Economic Forum, 2021).[2]

Besides fake reviews, manipulation may be less brazen, e.g., through *incentivized reviews*, where a customer is incentivized to provide a positive opinion; common examples of such incentives include entry into a sweepstake, coupon, or a discount. While incentivized reviews are banned on certain large platforms, e.g., Yelp and Amazon, other platforms allow for incentivized reviews (Techcrunch, 2017; Yelp, 2017; Federal Trade Commission, 2022). In other cases, manipulation may be completely innocuous, e.g., by providing additional after-sale services and care. For example, a seller of phone cases helped customers who ordered an incorrect product to fix the problem proactively, triggering a high-rating review from the customer (Figure 1a). In another example, a seller offered a full refund without requiring product return to a buyer posting a quality complaint, and the buyer subsequently revised the poor review voluntarily without seller request (Figure 1b).



Very nice case and service

Reviewed in the United States on November 25, 2022

Color: Black | Verified Purchase

This case was very durable. When buying it, I thought it would more of a silicone feel to it, but surprisingly it does not. It gives more of a sturdy feel to it. The one thing was that I accidentally bought the wrong case, and the seller was able to fix the problem and send me the right case. They had sent an email asking if the case was alright and made sure that I bought the right one. Communication was great. Since there aren't many cases for this type of phone, I am extremely satisfied with the case and the service that came with it!

(a)

Warranty provided by Seller is top notch

Reviewed in the United States on January 31, 2023

Color: Blue | Verified Purchase

Love this at first. Bought it on December 6th. It's an awesome case. But the feature that I liked, the circular stand in the back is broken. The stand no longer holds up. It immediately closes. Since it's past the 30 days for an Amazon return, I tried to find the warranty info. The warranty is 90 days. Online it says to contact us via amazon. Can't find out how to contact the seller... Asked the question in the review section and no one knows. So if and when the seller stands behind their product I might change my review.

EDIT 2/4/2023 - Once I posted this review I was contacted by the seller and offered a full refund with no need to return the item. I responded with acceptance of their offer and received the refund within 48 hours of my answer.

In addition, I WAS NOT asked to update my review. I'm doing so because the seller made it right and the product does protect my phone. I am ordering a different case from the same seller in hopes that the ring stand will last.

(b)

Figure 1. High-Rating Review for After Sale Care
Source: https://www.amazon.com

Irrespective of the nature of the manipulation, in all cases, sellers find it costly to manipulate consumer opinion. In addition, sellers may face platform filters or sanctions that further drive up the cost of manipulation. For example, Yelp uses automated software tools to identify and

[2]These estimates are based on self-reported data from Trip Advisor, Yelp, TrustPilot and Amazon (World Economic Forum, 2021).

remove reviews that are suspected to have been solicited (Yelp, 2017). Similarly, Amazon uses technology to identify and delete consumer opinions that are deemed fraudulent on its platform (He et al., 2022).

In this paper, we study the competitive landscape for online sellers that sell differentiated yet substitutable products on a platform. Sellers simultaneously determine their product price and their manipulation strategy, characterized by the extent to which sellers artificially inflate consumer opinion. Each consumer chooses the product that yields the highest utility among the available options. Consumer utility depends on product price and the *perceived* product quality – that consists of the *true* quality from the product's features, and the extent to which the seller inflates consumer opinion. There are competing arguments relating to which sellers have a greater incentive to manipulate consumer opinion. Dellarocas (2006) considers a market where a seller signals their true quality to uninformed consumers via manipulation. They show that manipulation is increasing in the true quality of the seller if the marginal benefit from higher perceived quality is increasing in its true quality. That is, if sellers stand to gain more from being perceived as high quality, then higher quality sellers manipulate more. In contrast, He et al. (2022) find that manipulation is predominantly employed by lower quality sellers. They argue that, while sellers of all qualities benefit from manipulation, the higher quality sellers find it a lot harder to manipulate, as opposed to the lower quality sellers. In this paper, we examine the equilibrium pricing and manipulation strategy of the sellers in an oligopoly. In particular, when sellers are heterogeneous in their true qualities, how does the equilibrium price and manipulation effort vary based on the true product quality? In light of the contradicting findings of Dellarocas (2006) and He et al. (2022), we explore the dynamics that drive sellers' manipulation incentives, both in their tendency to manipulate and in the extent of manipulation.

We also analyze how manipulation affects the platform and the consumers. Sellers' decision to manipulate consumer opinion affects their perceived product quality and their subsequent sales. The equilibrium product prices and manipulation effort affect the sellers' revenue. As a result, a platform which charges a commission on each transaction would see an impact on its revenue. In addition, the platform may be in a position to implement practices that affect sellers' ability to manipulate consumer opinion, e.g., by affecting the sellers' cost of manipulation. How should the platform exercise its leverage to optimize the platform revenue?

Furthermore, in the presence of manipulation, consumers' true utility from a product (ex-post purchase) differs from that drives their purchase decision (ex-ante expectation). Are consumers always losing in the manipulation game? In this paper, we build a model that encompasses these considerations and conduct equilibrium analysis and optimization to derive insights on seller competition, effect of review manipulation, and platform policy on customer reviews.

Before presenting our work, we emphasize that studying manipulation does not imply that we consent to or endorse such practices. Rather, we believe understanding its impact and the mechanism of its effect helps platforms and policy makers design strategies and policies that are effective in curbing unethical practices, reduce trade frictions, and improve market efficiency. Indeed, there has been significant interest in understanding the effects of manipulation by practitioners (The Wall Street Journal, 2023) and regulatory agencies, e.g., the FTC (Federal Trade Commission, 2023). Our work in this paper, on understanding the effects of manipulation on competition, is crucial to policymakers and managers.

2 Related Literature

This paper is closely related to two streams of literature: (a) models of competition using the MNL choice model and (b) empirical and theoretical models on firms' manipulation of customer opinion.

2.1 A Brief Background on The MNL Choice Model

Discrete choice models are widely used in Economics, Marketing, and OM to describe and analyze how consumers choose among a collection of alternatives. These models assume that consumers are random utility maximizers. The simplest and most studied discrete choice model is the multinomial logit (MNL) model (McFadden et al., 1973; Berry, 1994). Arguably, one of the most attractive features of the MNL model is in its empirical support to estimate model parameters with data. In their pioneering work, McFadden et al. (1973) establish the concavity of the log-likelihood function in the model parameters. Vulcano et al. (2012) propose an expectation-maximization (EM) algorithm to incorporate incomplete data (e.g., the "no-purchase" option) with the MNL model. We borrow these techniques in estimating our consumer choice model in Section 6.1 to conduct numerical experiments.

2.2 Models of Price Competition under The MNL Choice Model

The MNL model has been extensively employed for understanding firms' pricing decisions in oligopolistic competition in an economy. Due to the extensive nature of this stream, we mention papers within OM that are closely related to our work. One of the earliest papers in this stream is Anderson and De Palma (1992). They show the existence of an equilibrium when symmetric multiproduct firms compete in prices under the MNL demand and conclude that when all products have equal quality, the equilibrium prices are a fixed markup over the production cost. Besanko et al. (1998) and Besanko et al. (2003) propose a framework to empirically estimate logit demand systems where prices are assumed to be the equilibrium outcomes of Nash competition among manufacturers and retailers. Their work explains the bias that arises in model estimates when the endogeneity of prices is ignored. Earlier work by Berry et al. (1995) and subsequent work by Berry et al. (2004) empirically analyze the equilibrium prices under oligopolisitic competition in the US auto industry and obtain estimates of demand and cost parameters. The existence of a unique Nash equilibrium for price competition under the MNL model is established; see Gallego et al. (2006), Bernstein and Federgruen (2004) and Allon et al. (2011). Farahat and Perakis (2011) study models of competition for differentiated products, where firms compete either in prices (Bertrand) or quantity (Cournot), and demand follows the MNL model. They show that the outcomes under Bertrand and Cournot competition are respectively equivalent to outcomes when decisions are made sequentially: the Cournot outcome arises when the production decision precedes the pricing decision, while the Bertrand outcome arises when the pricing decision precedes the production decision. Li and Huh (2011) extend these models of competition to the case of the nested logit model and provide quasi-closed form expressions for the equilibrium market share and markups of firms. Gallego and Wang (2014) identify conditions that ensure a unique equilibrium under the nested logit model with product-specific price sensitivities. Aksoy-Pierson et al. (2013) and Lee and Çakanyildirim (2021) study price competition under the mixed MNL model and identify conditions for a unique Nash equilibrium. In this paper, we build upon this literature, particularly, that of Li and Huh (2011) to analyze the competition of multiple sellers on a common platform and study the effect of review manipulation in this setting.

Recently, Wang et al. (2022) analyze a model of competition under consumer choice models where firms compete in prices, quality and associated service duration (e.g., maintenance and warranty), where the associated service cost depends on the product quality: service costs are lower if product quality is higher. Our work differs from theirs in that we analyze the types of firms who manipulate consumer opinion and the effect of such manipulation on sellers, the platform and consumers under the equilibrium manipulation and pricing strategy of firms.

2.3 Empirical and Theoretical Models on Manipulation of Consumer Opinion

Since the dawn of e-commerce, one of the most important roles of platforms that match buyers and sellers has been the provision of information about products and sellers via consumer opinion/feedback, typically absent in offline environments. Such consumer opinion arises via ratings and review comments that are viewed by subsequent shoppers and influences their purchasing decisions. For example, Chevalier and Mayzlin (2006) and Luca (2011) quantify the marginal benefit from an increase in review rating in the context of books and restaurants, respectively. Beyond influencing other shoppers' purchasing behavior, consumer opinion plays an important role in platform's listing strategy (Chen and Wilson, 2017). As a result, sellers may intentionally manipulate their ratings in order to be perceived more attractive to entice more consumers. One of the earliest papers in this stream, Dellarocas (2003), discuss the challenges and opportunities brought by such feedback mechanisms.

Theoretical work in this stream spans multiple disciplines including OM, Information Systems (IS), Marketing, and Economics. We discuss papers closest to our work. Dellarocas (2006) analyze a market where a seller signals its quality via manipulation. Consumers update their beliefs on the seller's true quality based on observed signal (the sum of the true quality, the extent of manipulation, and a noise term). They show that the extent of manipulation depends on the marginal benefit from quality: If the marginal benefit is increasing in quality, then the extent of manipulation is increasing in true quality. Mayzlin (2006) analyzes an environment where sellers use promotional chat and consumers learn about the seller's quality. They show that in equilibrium, sellers with inferior products spend more resources purchasing promotional reviews. Relatedly, Sun (2012) analyzes the effect of variance in product ratings, and posits that a higher average rating corresponds to a higher quality, while a higher variance corresponds to a niche product (i.e., extreme in fit).

Empirically, Luca and Zervas (2016) and He et al. (2022) test the economic incentives for firms to purchase faudulant reviews and show the presence of a large and active market for manipulation. Luca and Zervas (2016) show that a restaurant on Yelp is more likely to manipulate if its reputation is weak. Further, restaurants are more likely to manipulate when the intensity of competition is strong. He et al. (2022) reach a similar finding that low quality sellers on Amazon are more likely to manipulate. Using controlled experiments, Ananthakrishnan et al. (2020) analyze a platform's information display strategy when it contains a mix of true and fraudulent reviews. They find that consumer trust is higher when the platform displays both true and fraudulent reviews instead of fully censoring fraudulent reviews.

The main results from extant literature show the polarity in the types of sellers engaged in consumer opinion manipulation. That is, either high-quality firms or low-quality firms choose to manipulate, which appears to be contradicting. In addition, little is known about how firms in the middle react. In this paper, we analyze how sellers' decisions to manipulate affects their prices under multi-seller competition when demand follows the MNL model. First, we identify a unique Nash equilibrium in an oligopoly, deriving the (quasi) closed-form expressions for the equilibrium markup and manipulation level. We then identify an index to measure a seller's propensity to manipulate and show that firms' propensity to manipulate may be increasing, decreasing, or unimodal in firms' true quality, depending on the cumulative volume of true reviews. These results identify a contiguous set of firms that choose to manipulate in equilibrium. Eventually, we investigate the conditions where firms and consumers may gain benefits or be hurt by review manipulation. We apply our theoretical results using data from Wang et al. (2014) to better understand firms' manipulation strategy.

3 Model

Consider a platform-enabled marketplace, consisting of n competing sellers, indexed by $i \in [n]$,[3] and a mass of potential consumers, normalized to 1. Each seller markets and sells a product with true quality a_i, unit cost c_i, and chooses price p_i.[4] A representative

[3] We denote the set $\{1, 2, \ldots, n\}$ by $[n]$.

[4] We refer to the product of seller i by product i. We use the terms *seller* and *firm* interchangeably.

consumer purchases exactly one product from the $[n] \cup \{0\}$ products, where 0 represents the no-purchase option. The consumer utility from purchasing product i depends on the following: the *perceived* quality of the product – the sum of *true* product quality a_i and the extent of *manipulation* x_i – and the price. Specifically, the (perceived) consumption utility from product i is as follows:

$$u_i = a_i + x_i - bp_i + \epsilon_i \tag{1}$$

where ϵ_i is an i.i.d standard Gumbel random variable and b is a price sensitivity parameter. We normalize $\mathbb{E}[u_0] = 0$. The market-share of seller i, denoted by q_i, follows from the standard MNL model and is shown below.

$$q_i = \frac{e^{u_i}}{1 + \sum_{j \in [n]} e^{u_j}} \ . \tag{2}$$

Denote the profit margin (*markup*) of product i by m_i, where $m_i = p_i - c_i$. For mathematical simplicity, define A_i as $A_i = e^{a_i - bc_i}$. We refer to A_i – the cost-adjusted quality of seller i – as the *type* of seller i, and refer to a seller with a higher value of A_i as a higher type. We rewrite seller i's market-share as follows:

$$q_i = \frac{e^{a_i + x_i - b(m_i + c_i)}}{1 + \sum_{j \in [n]} e^{a_j + x_j - b(m_j + c_j)}} = \frac{A_i e^{x_i - bm_i}}{1 + \sum_{j \in [n]} A_j e^{x_j - bm_j}} \tag{3}$$

Let $h_i(x_i)$ denote the cost of manipulation for seller i. The profit of seller i, denoted by π_i, is as follows:

$$\pi_i = \underbrace{(p_i - c_i)q_i}_{\substack{\text{Profit from} \\ \text{Direct Sales}}} - \underbrace{h_i(x_i)}_{\substack{\text{Cost of} \\ \text{Manipulation}}} = m_i \left(\underbrace{\frac{A_i e^{x_i - bm_i}}{1 + \sum_{j \in [n]} A_j e^{x_j - bm_j}}}_{=q_i} \right) - h_i(x_i). \tag{4}$$

To begin with, we analyze the outcome in the absence of any seller-manipulation and use this as a benchmark. Subsequently, we analyze the outcomes in the presence of seller-manipulation. We assume that the customers' price sensitivity, each seller's price and perceived quality are common knowledge to all sellers in the oligopoly. The unit production cost c_i and manipulation x_i do not need to be observable (public information), as the market share of each seller depends on the observed price (i.e., the sum $(m_i + c_i)$) and the perceived quality (i.e., the sum $(a_i + x_i)$) of other sellers but not their cost and manipulation directly. Each seller responds to the observed prices and perceived quality of other sellers by choosing its markup and manipulation to maximizes its own profit; once no seller can increase their profit by unilaterally deviating, an equilibrium is reached.[5]

4 Absence of Seller Manipulation

Suppose that manipulation is prohibitively expensive; thus, the sellers do not manipulate their perceived quality, i.e., $\mathbf{x} = \mathbf{0}$.[6] We denote this setting by AM (absence of manipulation). It has been established in the literature that a unique price equilibrium exists (e.g., Li and Huh 2011). We reproduce the results in this section to form a benchmark. Seller i's market share in (2) simplifies to:

$$q_i = \frac{A_i e^{-bm_i}}{1 + \sum_{j \in [n]} A_j e^{-bm_j}} \ . \tag{5}$$

We analyze seller i's best response enroute to identifying the equilibrium markups. Seller i's profit is

$$\pi_i = m_i q_i = m_i \left(\frac{A_i e^{-bm_i}}{1 + \sum_{j \in [n]} A_j e^{-bm_j}} \right) \ .$$

Fix $\mathbf{m}_{-i}$. Observe that π_i is unimodal in m_i. Seller i's best-response satisfies:

$$m_i(\mathbf{m}_{-i}) = \frac{1}{b(1 - q_i(m_i(\mathbf{m}_{-i}), \mathbf{m}_{-i}))}. \tag{6}$$

That is, $m_i(\mathbf{m}_{-i})$ is the unique solution to the following univariate equation in terms of m_i:

$$m_i = \frac{1}{b} \left(1 + \frac{A_i e^{-bm_i}}{1 + \sum_{j \neq i, j \in [n]} A_j e^{-bm_j}} \right).$$

Define $f(z)$ as follows:

$$f(z) \triangleq z e^{\frac{1}{1-z}}. \tag{7}$$

$f(x)$ is increasing in x, $f(0) = 0$ and $f(1) = \infty$. The following result identifies the equilibrium markups and the resulting market share of each seller.

Theorem 1 (Equilibrium Outcome under AM)**.** *The equilibrium q_0^{AM} is the solution to the following equation:*

$$q_0 = 1 - \sum_{j \in [n]} f^{-1}(A_i q_0) \tag{8}$$

The equilibrium market-share and markup of seller i is:

$$q_i^{\mathsf{AM}} = f^{-1}(A_i q_0^{\mathsf{AM}}) \text{ and } m_i^{\mathsf{AM}} = \frac{1}{b(1 - q_i^{\mathsf{AM}})}. \tag{9}$$

[5] Essentially, each seller responds to the observed values $p_i = m_i + c_i$ and $\hat{a}_i = a_i + x_i$ of other sellers without directly observing m_i and x_i despite that q_i is written as a function of the markup vector and manipulation vector.

[6] For any quantity of interest y, we denote $(y_j)_{j \in [n]}$ by $\mathbf{y}$, and $(y_j)_{j \in [n] \setminus \{i\}}$ by $\mathbf{y}_{-i}$.

From Theorem 1, we have that in the absence of manipulation, a seller with a higher type has a higher equilibrium margin, market share and profit.

5 Presence of Seller Manipulation (PM)

The result in Theorem 1 aligns well with the typical observation that stronger sellers do well in a competitive market. Now, suppose that sellers can manipulate their perceived quality. Are the high-type sellers less inclined to engage in quality manipulation because they are doing well, or are they compelled to dominate the market even more when given the chance to further elevate the market's perception of their quality, barring legal and moral obstacles? Recall from the discussion in Section 2.3 that the findings in the extant literature has been limited but mixed. Dellarocas (2006) argue that higher quality sellers have a greater incentive to manipulate while He et al. (2022) show empirical evidence that manipulation is predominantly employed by lower quality sellers. While these insights are derived either from a stylized theoretical setting or obtained from evidence in a particular data set, we examine the same question by evaluating a multi-seller price competition under the empirically-supported MNL demand model. We present a measure called "propensity to manipulate" that identifies sellers more inclined to manipulate, which unifies the theoretical and empirical observations in the literature.

Formally, seller i whose true quality is a_i manipulates their perceived quality to be $a_i + x_i$. We denote this setting by PM (presence of manipulation). Recall that x_i denotes the extent of manipulation by seller i, and $h_i(x_i)$ denotes the cost of manipulation. We make the following assumption on the cost of manipulation.

Assumption 1 (Cost of Manipulation)**.** *The cost of manipulation $h_i(x)$, $i \in [n]$ satisfies the following:*

(a) $h_i(x)$ *is smooth, non-negative, increasing and strictly convex in* $x \in \Re^+$, *i.e.,* $h_i(x) \geq 0, h_i'(x) \geq 0, h_i''(x) > 0$ *for* $x \geq 0$ *with* $h_i(0) = 0$.

(b) $h_i''(x) \geq h'(x)$ *for all* $x \in \Re^+$.

Part (a) is straightforward and assumes that it becomes increasingly more difficult to manipulate. Part (b) states that the cost function is *sufficiently* convex, a regularity condition that ensures π_i is well-behaved, i.e., it excludes irregularities in manipulation cost that may lead to multiple equilibria. In particular, observe that part (b) can be written as $h_i(x) \geq h_i'(0)\,(e^x - 1)$. While part (b) might appear restrictive at first sight, we note that this assumption applies to cost as a function of the *resulting increment* in perceived quality x, not that of the manipulation input. In Section 6, we will show that a quadratic cost function in terms of the number of solicited reviews (i.e., when a seller faces linear marginal cost to solicit fake reviews) satisfies the above assumption.

Below, we analyze seller i's best-response. Subsequently, we analyze the equilibrium of this game.

5.1 Best Response of Seller i

When a seller begins to manipulate, their perceived quality increases, affecting customer choice and disturbing any market equilibrium. Others will respond and their response is two-pronged - they may manipulate their own perceived quality and/or they may adjust their prices. These actions will in turn trigger a new round of responses until an equilibrium, if one exists, is reached. Consider seller i. Fix the decisions of all sellers other than i, i.e., $(\mathbf{m}_{-i}, \mathbf{x}_{-i})$. We analyze seller i's best response (m_i, x_i). We proceed by first deriving the optimal markup m_i as a function of any given x_i (Section 5.1.1), and then solving the optimal x_i (Section 5.1.2).

5.1.1 Optimal Markup m_i

Fix x_i. The following result identifies the optimal m_i.

Lemma 1. *Fix $(\mathbf{m}_{-i}, \mathbf{x})$. π_i is unimodal in m_i. Seller i's best-response $m_i(x_i; \mathbf{x}_{-i}, \mathbf{m}_{-i})$ satisfies $m_i = 1/b(1-q_i)$, where we omit the arguments of m_i and q_i for brevity. That is, $m_i(x_i; \mathbf{x}_{-i}, \mathbf{m}_{-i})$ is the unique solution to the following univariate equation of m_i:*

$$m_i = \frac{1}{b}\left(1 + \frac{A_i e^{x_i - bm_i}}{1 + \sum_{j \neq i} A_j e^{x_j - bm_j}}\right). \tag{10}$$

In addition, $m_i(x_i; \mathbf{x}_{-i}, \mathbf{m}_{-i})$ is increasing in x_i, increasing in $\mathbf{m}_{-i}$ and decreasing in $\mathbf{x}_{-i}$. Further, $x_i - bm_i(x_i; \mathbf{x}_{-i}, \mathbf{m}_{-i})$ is increasing in x_i.

Lemma 1 concludes that, fixing other sellers' decisions, a seller would adjust its manipulation and markup in tandem - a higher (lower) manipulation is matched with a higher (lower) markup. Although the increase in manipulation and markup has opposing effect – the former makes the seller's product more attractive (due to higher perceived quality) whereas the latter makes it less appealing (due to higher price) – the overall effect on product attractiveness, reflected through the term $x_i - bm_i$, is dominated by extent of manipulation.

For convenience, we denote $m_i(x_i; \mathbf{x}_{-i}, \mathbf{m}_{-i})$ with $m_i(x_i)$ and $q_i(x_i, m_i(x_i))$ with $q_i(x_i)$. From Lemma 1, all else equal, $q_i(x_i)$ is increasing in x_i.

5.1.2 Optimal Manipulation x_i

Fix $\mathbf{m}_{-i}, \mathbf{x}_{-i}$. For any choice x_i of seller i, their choice of m_i is given by (10) in Lemma 1. Below, we identify seller i's optimal choice of x_i (i.e., their best-response to $\mathbf{m}_{-i}, \mathbf{x}_{-i}$).

Lemma 2. *Fix $\mathbf{m}_{-i}, \mathbf{x}_{-i}$. $\pi_i(x_i)$ is quasi-concave in x_i. The optimal x_i^* is as follows:*

(a) If $\frac{q_i(0)}{b} \leq h_i'(0)$, then, $x_i^* = 0$.

(b) Otherwise, x_i^* is the solution to the following equation:

$$\underbrace{\frac{q_i(x_i)}{b}}_{\text{Marginal Benefit from Manipulation}} = \underbrace{h_i'(x_i)}_{\text{Marginal Cost of Manipulation}} . \tag{11}$$

Part (a) of Lemma 2 shows the condition under which seller i chooses not to manipulate: If the marginal cost of manipulation at $x_i = 0$ is too high, then the benefit from manipulation does not offset the cost of manipulation. Otherwise, seller i manipulates by a positive amount. The extent of manipulation is identified in part (b). The LHS in (11) corresponds to the marginal benefit from manipulation, while the RHS corresponds to the marginal cost of manipulation. At optimality, the marginal benefit is equal to the marginal cost. While the LHS and RHS of (11) are both increasing in x_i, due to Assumption 1(b), we can show that $\pi_i(x_i)$ is quasi-concave in x_i and the solution is unique.

Lemma 2 reveals the tension between two competing forces: the cost and benefit from manipulation. The marginal benefit from manipulation for a seller is proportional to its market share. This seems to suggests that, at least in theory, a seller with a stronger product (hence a larger market share) stands to benefit more from manipulation, as argued by Dellarocas (2006). We will illustrate in later analysis that this insight is only a partial view both in theory and in reality.

5.2 Equilibrium Outcome

We aim to identify the distinction between sellers that do and do not manipulate in the equilibrium. We also examine how their pricing strategies differ depending on their manipulation strategy. To that end, we derive a seller's equilibrium pricing and manipulation strategy in a competitive market. Let $\mathbf{x}^{\mathsf{PM}}$ and $\mathbf{m}^{\mathsf{PM}}$ denote the equilibrium manipulation and markups, respectively. Let $\mathcal{X}$ denote the set of sellers that manipulate in equilibrium, i.e.,

$$\mathcal{X} = \{i : x_i^{\mathsf{PM}} > 0\}. \tag{12}$$

Therefore, the set $\mathcal{X}^{\mathsf{C}} = [n] \setminus \mathcal{X}$ consists of sellers that do not manipulate in equilibrium, i.e., $\mathcal{X}^{\mathsf{C}} = \{i : x_i^{\mathsf{PM}} = 0\}$.

If the marginal cost of manipulation $h_i'(0)$ for all $i \in [n]$ is too large, then no seller chooses to manipulate (i.e., the set $\mathcal{X}$ is empty), and the equilibrium outcome under PM is identical to that under AM. For a non-trivial outcome under PM, we make the following assumption.

Assumption 2. *There exists some seller $i \in [n]$ s.t.*

$$\underbrace{h_i'(0)}_{\substack{\textit{marginal cost of manipulation}\\ \textit{at } x_i = 0}} < \underbrace{\frac{q_i^{\mathsf{AM}}}{b}}_{\substack{\textit{marginal benefit from manipulation to seller } i\\ \textit{at } x_i=0,\, \mathbf{x}_{-i}=\mathbf{0}}}. \tag{13}$$

Observe that the LHS is the marginal cost of manipulation at $x_i = 0$, while the RHS is the marginal benefit from manipulation at $x_i = 0$, $\mathbf{x}_{-i} = \mathbf{0}$, where q_i^{AM} is the equilibrium market share in the absence of manipulation.

Assumption 2 implies that, there exists at least one seller i such that, if no other seller were to manipulate, seller i has a strict incentive to manipulate. Consequently, under this assumption, the absence of manipulation does not constitute an equilibrium outcome (i.e., the set of sellers that manipulate in equilibrium, $\mathcal{X}$, is non-empty).

For any $i \in [n]$ and $z \in [bh_i'(0), 1)$, define the following:

$$g_i(z) \triangleq z e^{\frac{1}{1-z} - h_i'^{-1}\left(\frac{z}{b}\right)} = f(z) e^{-h_i'^{-1}\left(\frac{z}{b}\right)}, \tag{14}$$

where $f(z)$ is as defined in (7). In Lemma B1 in Appendix B, we show that $g_i(z)$ is increasing in z. Next, define γ_i (an index for seller i) as follows:

$$\gamma_i = \begin{cases} \frac{A_i}{f\left(bh_i'(0)\right)}, & \text{if } h_i'(0) < \frac{1}{b}; \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

The following result identifies each seller's equilibrium manipulation and markup.

Theorem 2 (Equilibrium Outcome under PM)**.** *The equilibrium outcome is as follows:*

(a) The set $\mathcal{X}$ is upward-closed in γ_i.

(b) The equilibrium market share of no-purchase, q_0^{PM}, is the unique solution to the following equation:

$$q_0 = 1 - \sum_{i \in \mathcal{X}} g_i^{-1}(A_i q_0) - \sum_{i \in \mathcal{X}^{\mathsf{C}}} f^{-1}(A_i q_0). \tag{16}$$

The equilibrium market share, manipulation, and markup of seller i are:

$$q_i^{\mathsf{PM}} = \begin{cases} g_i^{-1}(A_i q_0^{\mathsf{PM}}), & \textit{if } i \in \mathcal{X}; \\ f^{-1}(A_i q_0^{\mathsf{PM}}), & \textit{if } i \in \mathcal{X}^{\mathsf{C}} \end{cases}, \quad x_i^{\mathsf{PM}} = \begin{cases} h_i'^{-1}\left(\frac{q_i^{\mathsf{PM}}}{b}\right), & \textit{if } i \in \mathcal{X}; \\ 0, & \textit{if } i \in \mathcal{X}^{\mathsf{C}} \end{cases}, \quad \textit{and } m_i^{\mathsf{PM}} = \frac{1}{b(1 - q_i^{\mathsf{PM}})}.$$

(c) In equilibrium, firm $i \in \mathcal{X}$ iff $\gamma_i > 1/q_0^{\mathsf{PM}}$.

Theorem 2 presents the unique equilibrium solution of the price-manipulation competition. To understand part (a), suppose that $\gamma_1 \leq \gamma_2 \leq \ldots \gamma_n$. Part (a) states that the the set of sellers that manipulate in equilibrium consists of an upward-closed subset of $[n]$. That is, $\mathcal{X} \in \{\{n\}, \{n-1, n\}, \ldots, \{1, 2, \ldots, n\}\}$. Since f and g_i are monotone functions, the fixed point equation (16) in part (b) is easily solved through a bisection search. The equilibrium solutions of manipulation and markups follow. In sum, the price-manipulation equilibrium is tractable and computationally efficient.

In the next result, we demonstrate how γ_i can be interpreted as a seller's relative propensity to manipulate, and help demonstrate its importance in understanding market outcomes.

5.3 Interpreting γ_i: Seller's Propensity to Manipulate

Recall the definition of γ_i that measures a seller's quality relative to their cost of manipulation. Using Theorem 2, the following result characterizes the set $\mathcal{X}$ of sellers that manipulate, using a threshold structure on γ_i.

Corollary 1 (γ_i Measures Seller i's Propensity to Manipulate). *Suppose $\gamma_1 \leq \gamma_2 \leq \ldots \gamma_n$. Suppose seller i manipulates in equilibrium. Then, all sellers with a higher propensity than seller i will also manipulate in equilibrium.*

Intuitively, γ_i is large if seller i's has a higher type, or if they find it easy to manipulate (i.e., the marginal cost of manipulation, $h_i'(0)$, is low). Seller i is more likely to manipulate if γ_i is high. Note that the index γ_i depends solely on exogenous model parameters and not on the equilibrium outcome. Therefore, it is easy to compute. Recall $A_i = e^{a_i - bc_i}$, $h_i'(0)$ is seller i's marginal manipulation cost at zero manipulation, and $f(z) = ze^{\frac{1}{1-z}}$ is a positive and increasing function on $[0, 1]$ (sellers with $bh_i'(0) \geq 1$ will never manipulate; see Assumption 2). The value of γ_i presents a unified measure to detect and compare seller i's propensity to manipulate against others, and helps resolve contradicting arguments in the literature regarding which sellers are more likely to manipulate, as we explain next.

5.4 Homogeneous Cost of Manipulation

Under a homogeneous cost of manipulation, say $h_i(x) = h(x)$ for all i, observe that the denominator in (15) is fixed for all i. To identify the equilibrium $\mathcal{X}$, from Lemma 1, it suffices to compare A_i. Accordingly, we have the following result.

Lemma 3 (Sellers with Higher Type Manipulate More). *Suppose that the cost of manipulation is identical for all sellers, i.e., $h_i(x) = h(x)$ for all i. Suppose $A_1 \leq A_2 \leq \ldots \leq A_n$. Then, $\gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_n$. In addition, for $i \in [n-1]$, we have the following: (i) $x_i^{\mathsf{PM}} \leq x_{i+1}^{\mathsf{PM}}$, (ii) $q_i^{\mathsf{PM}} \leq q_{i+1}^{\mathsf{PM}}$, and (iii) $m_i^{\mathsf{PM}} \leq m_{i+1}^{\mathsf{PM}}$. Further, $\pi_i^{\mathsf{PM}} \leq \pi_{i+1}^{\mathsf{PM}}$.*

Observe, from Lemma 3, that under a homogeneous cost of manipulation, a seller with a higher type is more inclined to manipulate, exerts greater effort in manipulation, has a higher profit margin, and a higher market share.

5.5 Homogeneous Types of Sellers

Under homogeneous seller types, say $A_i = A$ for all i, observe that the numerator in (15) is fixed for all i. To identify the equilibrium $\mathcal{X}$, it follows from Lemma 1 to compare the marginal cost of manipulation.

Lemma 4 (Sellers with Lower Manipulation Costs Manipulate More)**.** *Suppose that the types of sellers are identical, i.e., $A_i = A$ for all i. Suppose $h_1'(x) \geq h_2'(x) \geq \ldots \geq h_n'(x)$ for all $x \geq 0$. Then, $\gamma_1 \leq \gamma_2 \leq \ldots \leq \gamma_n$. In addition, for $i \in [n-1]$, we have the following: (i) $x_i^{\mathsf{PM}} \leq x_{i+1}^{\mathsf{PM}}$, (ii) $q_i^{\mathsf{PM}} \leq q_{i+1}^{\mathsf{PM}}$, (iii) $m_i^{\mathsf{PM}} \leq m_{i+1}^{\mathsf{PM}}$. Further, $\pi_i^{\mathsf{PM}} \leq \pi_{i+1}^{\mathsf{PM}}$.*

Lemmas 3 and 4 are special cases under which we glean, respectively, how seller quality and cost affect their equilibrium manipulation behavior. Specifically, Lemma 3 depicts a scenario in which we observe the effect described in Dellarocas (2006), namely, high quality sellers manipulate more. Lemma 4 brings the effect of manipulation cost into focus. While the above results are intuitive and capture the essence of certain dynamics, they are simplified special cases that do not reflect the full reality. In practice, sellers vary both in cost and in quality. More generally, their cost and quality also interact. That is, their cost of manipulation can depend on their true quality. In online retailing, for example, customer review ratings are usually capped at 5-star, and if high quality has already helped a seller achieve a high rating, further improvement through manipulation is increasingly difficult and expensive. Below, we consider a log-separable cost function that is quality-dependent.

5.6 Log-Separable Cost Functions

Suppose that the cost of manipulation depends on a firm's cost-adjusted quality A in the following log-separable form:

$$h(x; A) = \mathcal{H}(A)h(x), \tag{17}$$

where $h(\cdot)$ satisfies Assumption 1, and $\mathcal{H}(\cdot)$ is continuous and non-negative over $\Re^+$. Consequently, $h_i(x) = \mathcal{H}(A_i)h(x)$. Define the *type-elasticity* of the cost of manipulation as follows:

$$\varepsilon_A = \frac{\partial \log \mathcal{H}(A)}{\partial \log A}.$$

The type elasticity of the cost of manipulation determines the increase in the cost of manipulation with an increase in the type of the seller.

Lemma 5. *Suppose $A_1 \leq A_2 \leq \cdots \leq A_n$. Under the log-separable cost function in (17), we have the following distinct outcomes:*

(a) γ_i *is increasing in* A_i *iff the following condition holds:* $\varepsilon_A < \frac{1}{1+\frac{z}{(1-z)^2}}$, *where* $z = bh'(0)\mathcal{H}(A)$. *Consequently,* $\mathcal{X}$ *is upward-closed in* $[n]$ *iff the above condition holds.*

(b) γ_i *is decreasing in* A_i *if* $\varepsilon_A > 1$. *Consequently,* $\mathcal{X}$ *is downward-closed in* $[n]$.

The solution under log-separable cost function reveals the condition under which the insight "high quality seller manipulates more" holds or not. Without these restrictions, it is not clear which of the two competing effects dominate. While higher quality sellers gain more from manipulation, the type elasticity of the cost of manipulation, ε_A, determines the increase in the cost of manipulation with an increase in the seller's type. Part (a) of Lemma 5 shows that if ε_A is small, then sellers with higher types manipulate. Part (b) of Lemma 5 shows that if ε_A is large, then sellers with lower types manipulate.

Beyond the special cases above, γ_i may not be monotone in A_i. In such a case, it is unclear which sellers manipulate in equilibrium. In what follows, we describe a heterogeneous cost function that applies specifically to the star rating system frequently observed in retail platforms that provide ratings and reviews. We analyze the equilibrium outcome under PM. Subsequently, in Section 6.1, using a real-world data set, we demonstrate how the degree to which sellers manipulate (i.e., x_i) varies with their true quality (a_i).

6 An Illustration with Real-World Data: Seller Manipulation by Soliciting Fake Reviews

To illustrate our results, we examine an environment where sellers in an online platform manipulate their perceived value by soliciting fake reviews. Consider a star rating system employed by the platform on a scale of 0 to R, i.e., each seller is associated with a (true) star-rating between 0 and R, that the seller may manipulate by soliciting promotional (fake) reviews. Often, online star-rating systems adopt a *one-star* to *five-star* scale (e.g., Amazon) where the lowest star rating is one, not zero. In this case, we can map *five-star* to a value of $R = 4$ and *one-star* to the value 0 or use an alternative affine transformation, without loss of generality; similar technique applies to alternatively scaled, e.g., three-star or ten-star rating systems.

In this section, we begin with an analytical characterization of seller's propensity to manipulate under the star-rating system, and how it changes with the seller's true rating average and volume. Employing iso-γ curves, we then partition the cardinal space of rating average and volume into three regions – cost prohibitive, cost dominant, and benefit dominant – and show that the distribution of sellers among these regions determines the exhibited relationship between seller type and manipulation tendency. This analysis lays the foundation for our numerical experiments based on a real-world dataset in Section 6.1.

In the absence of any manipulation, let seller i's true rating be denoted by r_i^{tr} where $r_i^{\mathsf{tr}} \in [0, R]$. Let v_i^{tr} denote the volume of true ratings for seller i on the platform. Seller i manipulates their perceived rating to be higher than r_i^{tr}. Suppose the seller purchases v_i^{f} fake reviews with rating R.[7] Then, the observed rating for seller i is:

$$r_i^{\mathsf{ob}} \;=\; \frac{v_i^{\mathsf{tr}}}{v_i^{\mathsf{tr}}+v_i^{\mathsf{f}}} r_i^{\mathsf{tr}} + \frac{v_i^{\mathsf{f}}}{v_i^{\mathsf{f}}+v_i^{\mathsf{tr}}} R \;=\; r_i^{\mathsf{tr}} + \frac{v_i^{\mathsf{f}}}{v_i^{\mathsf{tr}}+v_i^{\mathsf{f}}}\left(R - r_i^{\mathsf{tr}}\right).$$

Empirically, we observe r_i^{ob} and $v_i^{\mathsf{tr}} + v_i^{\mathsf{f}}$. Let a consumer's utility from purchasing product i be denoted as follows:

$$\begin{aligned} u_i &= \beta_0 + \beta_r r_i^{\mathsf{ob}} + \beta_p p_i + \epsilon_i & (18)\\ &= \underbrace{\beta_0 + \beta_r r_i^{\mathsf{tr}}}_{a_i} + \underbrace{\beta_r \frac{v_i^{\mathsf{f}}}{v_i^{\mathsf{tr}}+v_i^{\mathsf{f}}}\left(R - r_i^{\mathsf{tr}}\right)}_{x_i} + \underbrace{\beta_p p_i}_{-bp_i} + \epsilon_i. & (19)\end{aligned}$$

Recall our consumer utility in (1); the quantities a_i, x_i and $-bp_i$ correspond to the quantities shown above. The seller's type corresponds to $A_i = e^{\beta_0+\beta_r r_i^{\mathsf{tr}}+\beta_p c_i}$. That is, given cost c_i, a seller's type only depends on the true rating r_i^{tr}. To purchase y fake reviews, let the cost incurred by a seller be the following:[8]

$$\text{Cost to purchase } y \text{ fake reviews} = k_1 y + k_2 y^2 \tag{20}$$

where $k_1, k_2 > 0$. Since seller i purchases v_i^{f} fake reviews, their cost of manipulation is $k_1 v_i^{\mathsf{f}} + k_2 {v_i^{\mathsf{f}}}^2$. By algebraic transformation of $x_i = \beta_r \frac{v_i^{\mathsf{f}}}{v_i^{\mathsf{tr}}+v_i^{\mathsf{f}}}\left(R - r_i^{\mathsf{tr}}\right)$ to express v_i^f and then substituting in (20), we have

$$h_i(x_i) = k_1\left(v_i^{\mathsf{tr}} \frac{\frac{x_i}{\beta_r}}{R - r_i^{\mathsf{tr}} - \frac{x_i}{\beta_r}}\right) + k_2\left(v_i^{\mathsf{tr}} \frac{\frac{x_i}{\beta_r}}{R - r_i^{\mathsf{tr}} - \frac{x_i}{\beta_r}}\right)^2. \tag{21}$$

The cost of manipulation in (21) satisfies Assumption 1 if k_2 is sufficiently larger than k_1 (see Appendix E for a detailed proof). Observe that the cost function above depends on both the true rating r_i^{tr} and the volume of true reviews v_i^{tr}.

For ease of notation, define the function $l(z)$ for $z \in (0,1)$: $l(z) = z + \left(\frac{z}{1-z}\right)^2$. We define two useful threshold values of v_i.

$$\overline{\overline{v}} = \frac{\beta_r R}{(-\beta_p)k_1}. \tag{22}$$

Let $\overline{v}$ denote the largest root of the following equation:

$$v = \overline{\overline{v}}\, l^{-1}\big((-\beta_p)k_1 v\big).$$

[7] We focus on manipulation by acquiring *additional* fake reviews instead of modifying existing poor reviews.

[8] While the cost to purchase y fake reviews is identical across sellers, the effect of y fake reviews is asymmetric across sellers. Specifically, from a fixed number of fake reviews purchased, seller i experiences a small (resp., large) x_i if v_i^{tr} is large (resp., small) or r_i^{tr} is large (resp., small).

It can be verified that if $\beta_r \leq \frac{1}{R}$, then $\overline{v} = 0$; otherwise, $\overline{v} > 0$. Since $l^{-1}(y) \in (0,1)$ for any $y > 0$, it follows that $\overline{v} < \overline{\overline{v}}$. Recall the definition of γ_i from (15), which corresponds to the index that measures seller i's propensity to manipulate. Using (15) and (21), we have:

$$\gamma_i = \begin{cases} \frac{\exp(\beta_0 + \beta_r r_i^{\mathsf{tr}} + \beta_p c_i)}{f\left(\frac{(-\beta_p)k_1 v_i^{\mathsf{tr}}}{\beta_r (R - r_i^{\mathsf{tr}})}\right)}, & \text{if } v_i^{\mathsf{tr}} + \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1} < \overline{\overline{v}}; \\ 0, & \text{otherwise.} \end{cases} \tag{23}$$

The condition $v_i^{\mathsf{tr}} + \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1} < \overline{\overline{v}}$ is equivalent to $h_i'(0) \leq \frac{1}{b}$. If this condition does not hold, then $\gamma_i = 0$ (from (15)). In particular, observe from (23) that if $v_i^{\mathsf{tr}} \geq \overline{\overline{v}}$, then $\gamma_i = 0$ regardless of the value of r_i^{tr}. The following result shows how seller i's propensity to manipulate (γ_i) changes with the seller's true rating (r_i^{tr}) and volume of ratings (v_i^{tr}).

Lemma 6. *(a) Fix r_i^{tr}. The propensity to manipulate γ_i is strictly decreasing in the volume of ratings v_i^{tr} and drops to 0 if $v_i^{\mathsf{tr}} > \overline{\overline{v}} - \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1}$.*

(b) Fix $v_i^{\mathsf{tr}} \leq \overline{\overline{v}}$.

(i) Suppose $v_i^{\mathsf{tr}} > \overline{v}$. Then, γ_i is decreasing in r_i^{tr} if $r_i^{\mathsf{tr}} \leq \frac{(-\beta_p)k_1}{\beta_r}\left(\overline{\overline{v}} - v_i^{\mathsf{tr}}\right)$ and drops to 0 if $r_i^{\mathsf{tr}} \geq \frac{(-\beta_p)k_1}{\beta_r}\left(\overline{\overline{v}} - v_i^{\mathsf{tr}}\right)$.

(ii) Suppose $v_i^{\mathsf{tr}} \leq \overline{v}$. Then, γ_i is unimodal in r_i^{tr}. That is, denote $\overline{r}$ as follows:

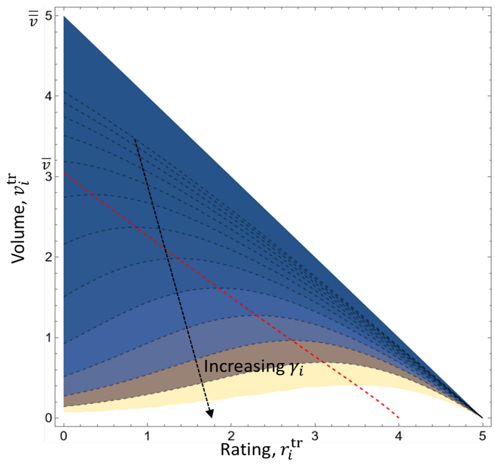
$$\overline{r} = \frac{(-\beta_p)k_1}{\beta_r}\left(\overline{\overline{v}} - \frac{v_i^{\mathsf{tr}}}{l^{-1}\left((-\beta_p)k_1 v_i^{\mathsf{tr}}\right)}\right) \tag{24}$$

γ_i is increasing in r_i^{tr} iff $r_i^{\mathsf{tr}} < \overline{r}$, is decreasing iff $\overline{r} < r_i^{\mathsf{tr}} < \frac{(-\beta_p)k_1}{\beta_r}\left(\overline{\overline{v}} - v_i^{\mathsf{tr}}\right)$, and drops to 0 if $r_i^{\mathsf{tr}} \geq \frac{(-\beta_p)k_1}{\beta_r}\left(\overline{\overline{v}} - v_i^{\mathsf{tr}}\right)$. Further, $\overline{r}$ is decreasing in v_i^{tr}.

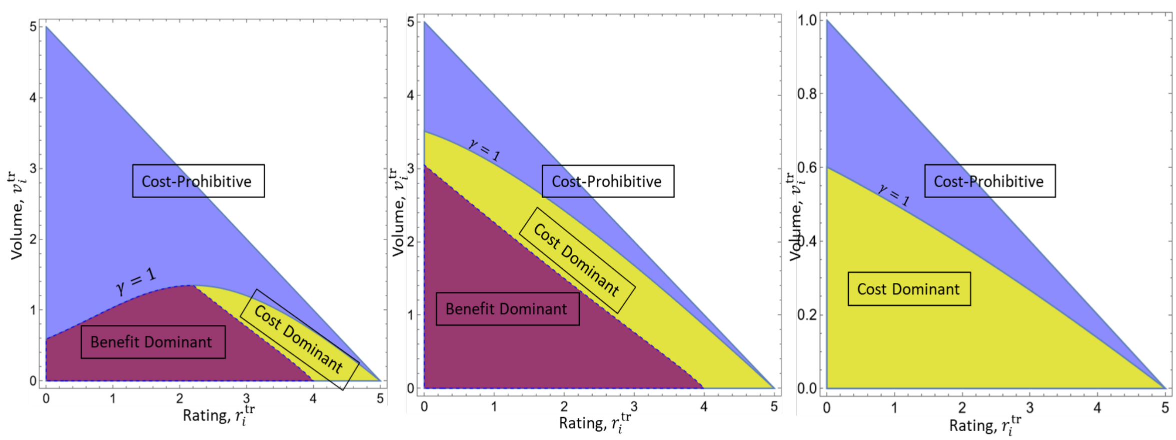
In Figure 2(a), we plot the iso-γ curves in the volume-rating coordinates using (23), with each point on a given contour corresponding to a fixed value of γ. The value of γ – the propensity to manipulate – increases in the direction marked by the arrow. As shown in part (b) of Lemma 6, $\overline{v}$ is the threshold value of v_i^{tr} that distinguishes two scenarios: (i) if $v_i^{tr} \geq \overline{v}$, then γ_i monotonically decreases in r_i^{tr}, which leads to the monotone iso-γ curves, and (ii) if $v_i^{tr} < \overline{v}$, then γ_i is unimodal in r_i^{tr}, which leads to the unimodal iso-γ curves (the dashed red curve that passes $(0, \overline{v})$ marks the mode of each iso-γ curve). The unshaded area bordered by the downward sloping line that passes $(0, \overline{\overline{v}})$ satisfies $v_i^{\mathsf{tr}} + \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1} \geq \overline{\overline{v}}$, hence $\gamma = 0$ in this region.

In Figure 2(b), we explain in detail the regions that arise in the rating-volume coordinates that distinguish seller behavior.

- Cost-Prohibitive Region: If either $\gamma_i = 0$, or $0 < \gamma_i \leq 1$, the marginal benefit from manipulation does not exceed the marginal cost at $x_i = 0$, and hence sellers do not



(a) Iso-γ contours in the $(r_i^{\mathsf{tr}}, v_i^{\mathsf{tr}})$ space.



(b) Regions that arise in the $(r_i^{\mathsf{tr}}, v_i^{\mathsf{tr}})$ space

Figure 2. Bottom: The white region corresponds to $\gamma_i = 0$, the blue region corresponds to $0 < \gamma_i < 1$ and the red and yellow region correspond to $\gamma_i > 1$. In particular, the red (resp., yellow) region corresponds to the case where γ_i is increasing (resp., decreasing) in r_i^{tr} for fixed v_i^{tr}. Values of parameters: Left: $-\beta_p = k_1 = c_i = 1$, $R = 5$, $\beta_r = 1$, $\beta_0 = 0$. Middle: $-\beta_p = k_1 = c_i = 1$, $R = 5$, $\beta_r = 1$, $\beta_0 = 3$. Right: $-\beta_p = k_1 = c_i = 1$, $R = 5$, $\beta_r = 0.2$, $\beta_0 = 3$.

manipulate. In Figure 2(b), the condition $\gamma_i = 0$ corresponds to the white region, while the condition $0 < \gamma_i \leq 1$ corresponds to the blue region. In both these regions, sellers do not manipulate since their cost of manipulation is large.

- Cost-Dominant Region: Recall the definition of $\bar{r}$ from (24) in Lemma 6(b)(ii). We refer to the region between $r_i^{\mathsf{tr}} \geq \bar{r}(v_i^{\mathsf{tr}})$ and $\gamma_i > 1$ as the cost-dominant region (the yellow region). In this region, a seller's propensity to manipulate decreases with its true quality.

- Benefit-Dominant Region: In the first figure of Figure 2(b), the region below $r_i^{\mathsf{tr}} \leq \overline{r}(v_i^{\mathsf{tr}})$ and $\gamma_i > 1$ is referred to as the benefit-dominant region (the red region). Here, a seller's propensity to manipulate increases in it's true quality. The iso-curve $\gamma = 1$ separates the cost-prohibitive region from the cost- and benefit-dominant regions. Notice that if $\overline{r}(\overline{v}) \leq 0$, then we obtain the second plot in Figure 2(b), in which the benefit-dominant region is not bordered by the iso-γ curve of $\gamma = 1$; if $\beta_r R < 1$, then $\overline{v} = 0$, and hence the benefit-dominant region disappears (the third plot in Figure 2(b)).

Fixing v_i^{tr}, in the benefit-dominant (resp., cost-dominant) region, γ_i increases (resp., decreases) in r_i^{tr}. In other words, seller i's propensity to manipulate increases with its true quality in the benefit-dominant region but decreases with its true quality in the cost-dominant region. What is the managerial interpretation of this observation? The effects of manipulation are manifested through the benefit from the manipulated increase in rating against the manipulation cost. All else equal, sellers with higher quality benefit more from manipulation but also incur higher manipulation cost; which of the two effects dominates depends on a seller's location in the volume-rating graph. Finally, we remark that, a seller located in the region $\gamma > 1$ does not necessarily manipulate in equilibrium. They only do so if γ_i exceeds $1/q_0^{\mathsf{PM}}$, which depends on the characteristics of all participating sellers.

One of the unique contributions of our paper is that a manager can identify these regions (i.e., the cost-prohibitive region, the cost-dominant region, and the benefit-dominant region), as well as the iso-γ curves in the volume-rating coordinates based solely on model parameters. For any given set of sellers in the competition, we can precisely locate each individual seller on the graph and identify the region it belongs to. This ability provides not only a managerial tool for understanding sellers' tendency to manipulate, but also instructive to a manager to monitor and predict how a seller's dynamically changing status of $(r_i^{\mathsf{tr}}, v_i^{\mathsf{tr}})$ shifts its propensity to manipulate as time evolves. The above can be accomplished without any equilibrium computation. To identify the exact set of sellers that manipulate in equilibrium, i.e., the set $\mathcal{X}$, we note that $\mathcal{X}$ is upward closed with respect to γ and satisfies $\gamma_i > 1/q_0^{\mathsf{PM}}$, where q_0^{PM} is obtained through equation (16) in Theorem 2.

Recall the contradicting findings of Dellarocas (2006) and He et al. (2022) on the relationship between quality and manipulation, the former suggesting that high quality (i.e., high r_i^{tr}) sellers are more likely to manipulate whereas the latter concluding that low quality (i.e., low r_i^{tr}) sellers are more likely to manipulate. The following corollary is a consequence of Lemma 6 and sheds light on resolving the contradicting views in the existing literature and provides a unified perspective.

Corollary 2. *Suppose Assumption 2 holds, $v_i^{\mathsf{tr}} = v$ for all $i \in [n]$, and $r_1^{\mathsf{tr}} \leq r_2^{\mathsf{tr}} \leq \dots r_n^{\mathsf{tr}}$. Then, $\mathcal{X}$ is contiguous. In addition,*

(a) The set $\mathcal{X}$ is downward-closed in r_i^{tr} if $r_1^{\mathsf{tr}} \geq \overline{r}(v)$.

(b) The set $\mathcal{X}$ is upward-closed in r_i^{tr} if $r_n^{\text{tr}} \leq \overline{r}(v)$.

Corollary 2 shows the various outcomes that emerge in equilibrium. On the one extreme, $\mathcal{X}$ can be downward closed, e.g., in markets for mature products. This conforms with the insights in He et al. (2022), who show that low quality sellers are likely to manipulate. On the other extreme, $\mathcal{X}$ can be upward closed, e.g., in nascent markets. This conforms with the predictions in Dellarocas (2006). In general, the set $\mathcal{X}$ may be neither upward- nor downward-closed; however, it is contiguous. Because all of the above scenarios are likely to occur in practice, one must take precaution in generalizing the observed trend. For example, when a dataset indicates a negative association between review ratings and manipulation, one cannot extend the association to sellers in other markets or even to extrapolate the trend to draw conclusions regarding other sellers in the same market. We illustrate this potential pitfall with Figure 3.

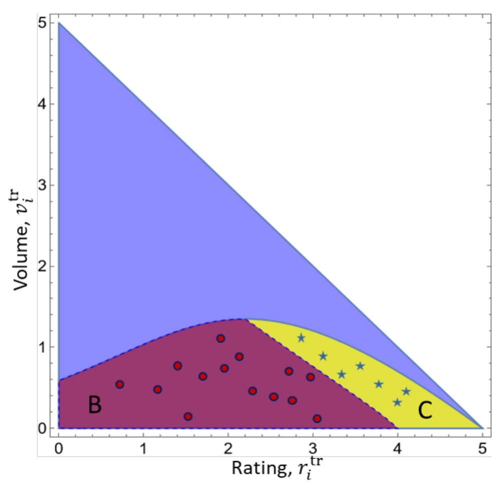


Figure 3. A snapshot of seller characteristics.

In Figure 3, sellers represented by red circles fall into the benefit-dominant region B and sellers marked with blue stars fall into the cost-dominant region C. Suppose the current dataset contains only red sellers in region B. Then the dataset may exhibit a negative relationship between quality and manipulation, while controlling for the volume of reviews. Likewise, if the available dataset contains only sellers in region C, it may reveal a positive connection between quality and manipulation. Clearly, extrapolating either exhibited trend to all sellers in the market can lead to misunderstanding of the market and ill-informed business decisions. Our findings caution decision makers against such fragmented views of the market. This insight applies not only to existing sellers in a market, but also future entrants to the market.

6.1 Numerical Application

To demonstrate the practical applications of our model, we assemble data from three datasets published by Wang et al. (2014) pertaining to electronic products, where the authors scrape amazon.com over a span a period of 24 weeks beginning February 1st, 2012. The first dataset comprises of transaction data for 2,163 unique products, the second dataset contains detailed product characteristics for 794 products, such as Operating System, RAM, processor, processor brand, storage size, average battery life (in hours), screen size, screen resolution, item weight, wireless type, mobile broadband, and webcam resolution and the third dataset comprises of customer reviews and provides information at the reviewer level, including review contents, post date, and review ratings.

6.1.1 Model Calibration

First, we select a set of $n = 11$ products that are close substitutes, with similar product features such as storage size and screen resolution. These products also have complete transactional information spanning the entire duration of $T = 24$ weeks. We infer the marginal production cost for each product from its selling price and the profit margin from the firms' financial statements if they are publicly listed, or the profit margin of their public competitors as a proxy for products sold by private firms. As Amazon provides information only on sales rank and not sales for each product in their data, we use a mapping from sales rank to sales rate to infer product sales in each time period. This approach has been widely employed in the literature, e.g., see Chevalier and Goolsbee (2003) and He et al. (2022). The mapping between sales s_{it} of product $i \in [n]$ in period $t \in [T]$ and its sales rank R_{it} is as follows:

$$s_{it} = e^{\frac{\beta}{\theta}} \frac{1}{(R_{it} - 1)^{\frac{1}{\theta}}} \tag{25}$$

Previous research, e.g., by Chevalier and Goolsbee (2003); He et al. (2022) report estimates of $\theta = 1.2$ and $\beta = 9.6$. We use these estimates to calculate s_{it} based on R_{it}. Then, customer utility u_{it} for product $i \in [n]$ in period $t \in [T]$ is modeled as shown in (18):

$$u_{it} = \beta_0 + \beta_r r_{it}^{\mathsf{ob}} + \beta_p p_{it} + \epsilon_{it},$$

where r_{it}^{ob} is the observed rating for product i at the start of period t, p_{it} is the price for product i in period t, and ϵ_{it} is i.i.d. Gumbel distributed. The price p_{it} is directly observed from the data, but r_{it}^{ob} is inferred as the average rating amongst all posted ratings until the start of period t. That is, we assume that the observed rating in period t is the average of the ratings "thus far" (i.e., until period $t-1$). Let $\mathfrak{r}_{it}$ denote the set of ratings posted for product i in period t. For any period $t \geq 2$, we calculate r_{it}^{ob} as follows:

$$r_{it}^{\mathsf{ob}} = \frac{\sum_{t' \in [t-1]} \sum_{r \in \mathfrak{r}_{it'}} r}{|\cup_{t' \in [t-1]} \mathfrak{r}_{it'}|}.$$

From the above consumer choice model, we have q_{it} (the probability that a representative consumer purchases product i at time t) as follows:

$$q_{it} = \frac{e^{\beta_0+\beta_p p_{it}+\beta_r r^{\mathsf{ob}}_{it}}}{1+\sum_{j\in[n]} e^{\beta_0+\beta_p p_{jt}+\beta_r r^{\mathsf{ob}}_{jt}}}. \tag{26}$$

Since our data does not contain information about no-purchase (s_{0t}), we employ the Expectation-Maximization (EM) algorithm (McLachlan and Krishnan, 2007) to estimate the parameters $\beta_0, \beta_p, \beta_r$. The results from our EM estimation procedure are as follows:

Table 1. Estimation Results

	β_p	β_r	β_0
Estimate	-0.0017***	0.103***	0.0109
Std. Error	0.000098	0.004	0.54

$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

6.1.2 Application and Results

Using the results from our estimation in Section 6.1.1 above and the analysis in Section 6, we predict the extent of manipulation in equilibrium for each of the n firms in our data. However, the dataset does not indicate whether each review is true or fake. We train a Long Short-Term Memory (LSTM) recurrent neural network model on a separate dataset provided by Salminen et al. (2022), where each review is labeled true or fake. While the nature of products across the two datasets is different, previous studies in Computer Science and Natural Language Processing report commonalities among fake reviews (Fang et al., 2020; Mohawesh et al., 2021).

From Table 1, $\beta_r = 0.103 < \frac{1}{R} = 0.25$. Therefore, in this application, the benefit-dominant region vanishes and sellers are located in either the cost-prohibitive region (if $0 \le \gamma \le 1$) or cost-dominant region. In Figure 4, we plot the regions in the volume-rating space and illustrate how the regions shift as the cost parameter changes. Recall that the regions $0 \le \gamma \le 1$ are cost-prohibitive and the region below the $\gamma = 1$ curve is the cost-dominant region. Sellers located below the $\gamma = 1/q_0^{\mathsf{PM}}$ curve (i.e., satisfy $\gamma > 1/q_0^{\mathsf{PM}}$) manipulate in equilibrium and those above the $\gamma = 1/q_0^{\mathsf{PM}}$ curve (i.e., satisfy $\gamma < 1/q_0^{\mathsf{PM}}$) do not manipulate. When k_1 is low ($k_1 = 0.1$; the left plot in Figure 4), all eleven sellers are in the cost-dominant region, and all but one manipulate in equilibrium. As k_1 increases ($k_1 = 1$; the middle plot in Figure 4), three sellers shift from the cost-dominant region to cost-prohibiting regions (one with $\gamma = 0$ and two with $\gamma \in (0, 1)$) and eight remains in the cost-dominant region, of which five manipulate in equilibrium; when k_1 increases to 3, six sellers shift to cost-prohibiting regions, and five remains in the cost-dominant region, of which only two manipulate. We remark that in these graphs, only the iso-γ curve $\gamma = 1/q_0^{\mathsf{PM}}$ requires equilibrium computation, whereas the rest are computed directly from model parameters. Therefore, the empirical versatility of the MNL model and our approach enable an easy-to-understand tool for market analysis that is both theoretically sound and managerially appealing, as illustrated in Figure 4.

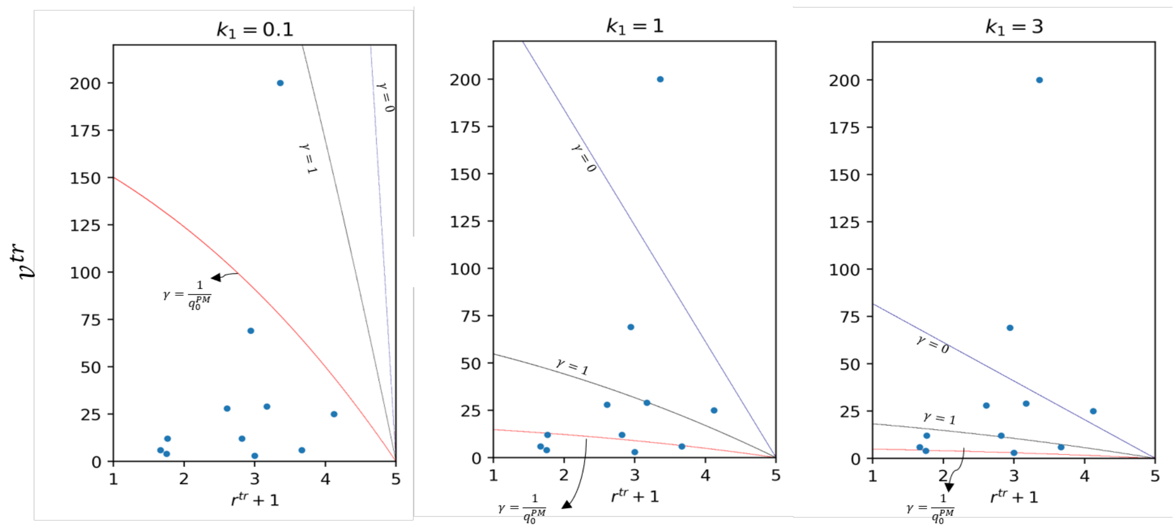


Figure 4. Seller Behavior with Changes in Cost of Manipulation.

Finally, the absence of the benefit-dominant region leads to the observation that, *ceteris paribus*, low quality sellers tend to manipulate, which ties back to the phenomenon emphasized in He et al. (2022). Our analysis indicates that, such phenomenon could occur for a specific market under certain conditions on model parameters but should not be viewed as universal truth.

7 Comparison of Equilibrium Outcomes

In an ideal world, sellers do not manipulate for ethical or legal considerations. In practice, such compliance is not guaranteed. Measures to prohibit manipulation and efforts to screen manipulated reviews are costly to platforms. To better understand the economic implications of manipulation, we compare the market outcomes in the absence and presence of seller manipulation. Several practical questions arise in this comparison. First, is manipulation profitable to sellers who choose to manipulate? Do these sellers see a higher market share by engaging in manipulation? Second, how does the presence of manipulation affect the industry and platform's revenues? Third, from a policy maker's standpoint, how does manipulation affect consumer welfare?

7.1 Effect of Manipulation on Market Shares, Profit Margins and Prices

Recall the definition of $\mathcal{X}$ (the set of sellers that manipulate in equilibrium) from (12). Define the set of sellers that have a higher market share and higher profit under PM as follows:

$$\mathcal{Q} = \{i : q_i^{\mathsf{PM}} \geq q_i^{\mathsf{AM}}\} \quad \text{and} \quad \Pi = \{i : \pi_i^{\mathsf{PM}} \geq \pi_i^{\mathsf{AM}}\}.$$

The following result characterizes the equilibrium market share of the no-purchase option.

Lemma 7. *The market share of the no-purchase option is lower in the presence of manipulation, i.e.,* $q_0^{\mathsf{PM}} < q_0^{\mathsf{AM}}$.

Consequently, $\sum_i q_i^{\mathsf{PM}} > \sum_i q_i^{\mathsf{AM}}$, i.e., the overall sales volume increases. Therefore, $\exists i$ s.t. $q_i^{\mathsf{PM}} > q_i^{\mathsf{AM}}$, i.e., at least one seller would sell more under PM than AM. Hence, $\mathcal{Q}$ is non-empty. For analytical tractability, in the remainder, we assume that sellers are homogeneous in their cost of manipulation. The next result confirms our intuition that sellers who do not manipulate would suffer in market share (relative to that under AM). However, it does not ensure that the converse or the inverse holds true. That is, sellers who manipulate do not necessarily gain market share. Further, recall that $m_i = \frac{1}{b(1-q_i)}$ holds in equilibrium under both AM and PM (Theorem 1 and Lemma 1); this also means that sellers who manipulate do not necessarily have a higher profit.

Theorem 3. *Suppose* $A_1 \leq A_2 \leq \ldots \leq A_n$ *and sellers are homogeneous in their cost of manipulation. The following statements hold:*

(a) Suppose $i \in \mathcal{X}^{\mathsf{C}}$. *Then,* $q_i^{\mathsf{PM}} < q_i^{\mathsf{AM}}$ *(i.e.,* $i \in \mathcal{Q}^{\mathsf{C}}$*) and* $m_i^{\mathsf{PM}} < m_i^{\mathsf{AM}}$. *Consequently,* $\pi_i^{\mathsf{PM}} < \pi_i^{\mathsf{AM}}$. *Stated differently,* $\mathcal{X}^{\mathsf{C}} \subseteq \mathcal{Q}^{\mathsf{C}} \subseteq \Pi^{\mathsf{C}}$, *or equivalently,* $\Pi \subseteq \mathcal{Q} \subseteq \mathcal{X}$.

(b) The sets $\mathcal{Q}$ *is upward-closed in* $[n]$.

Part (a) shows that if seller i does not manipulate, then their market share is strictly lower, i.e., $\mathcal{X}^{\mathsf{C}} \subseteq \mathcal{Q}^{\mathsf{C}}$. Since the equilibrium markup and market share move in tandem ($m_i = \frac{1}{b(1-q_i)}$), seller i's profit is also strictly lower. Part (a) can also be stated as $\Pi \subseteq \mathcal{Q} \subseteq \mathcal{X}$. An implication from this result is the necessity for seller i to manipulate in equilibrium so that their profit under PM exceeds that under AM. Nevertheless, it does not guarantee that manipulation increases seller i's profit (relative to AM). That is, it remains to be seen whether sellers that manipulate can even lose market-share and have lower profits under PM.

Part (b) states that the set $\mathcal{Q}$ is of the form $\{j^*, j^*+1, \ldots, n\}$ for some $j^* \in [n]$. From Lemma 7, we have that $\mathcal{Q}$ is non-empty. Together, we have that seller $n \in \mathcal{Q}$. This reinforces the insight that, under homogeneous cost, sellers with higher quality has a higher tendency to manipulate. However, the same cannot be said about Π. In what follows, we show that the ability to manipulate may hurt all sellers in equilibrium, i.e., it is possible that $\Pi = \emptyset$, using a simple example.

7.2 Can Manipulation Hurt All Sellers?

Consider the homogeneous cost function $h(x) = \lambda(e^x - 1)$, where $\lambda > 0$ is a cost-multiplier. Assumptions 1 and 2 imply that $\lambda < \frac{q_n^{\mathsf{AM}}}{b}$. We analyze the impact of an increase in the cost of manipulation – specifically, λ – on the market outcome.

Theorem 4. *Suppose $A_1 \leq A_2 \leq \ldots \leq A_n$. The following statements show the effect of an increase in λ on the equilibrium market share and profits of each seller.*

(a) *Consider seller $i \in \mathcal{X}$ (resp., $i \in \mathcal{X}^{\mathsf{C}}$). The equilibrium market share of seller i is decreasing (resp., increasing) in λ.*

(b) *Define τ as $\tau = \left(\sum_{i\in\mathcal{X}} \frac{A_i}{g'(q_i^{\mathsf{PM}})}\right) / \left(1 + \sum_{i\in\mathcal{X}} \frac{A_i}{g'(q_i^{\mathsf{PM}})} + \sum_{i\in\mathcal{X}^{\mathsf{C}}} \frac{A_i}{f'(q_i^{\mathsf{PM}})}\right)$. Suppose the following condition holds:*

$$(1-\tau)(2-q_n^{\mathsf{PM}})q_n^{\mathsf{PM}}/b < \lambda < q_n^{\mathsf{PM}}/b. \tag{27}$$

Then, for any $i \in [n]$, seller i's profit is increasing in λ.

Part (a) of Theorem 4 shows the effect of an increase in the cost of manipulation on a seller's equilibrium market share. An increase in the cost of manipulation results in a higher market share for sellers that do not manipulate, and a lower market share for sellers that manipulate. More importantly, part (b) of Theorem 4 shows the effect an increase in the cost of manipulation can have on a seller's equilibrium profit. In particular, the profit of *all* sellers increases as it becomes harder to manipulate, i.e., as the cost of manipulation increases. Stated differently, *all sellers benefit* from an increase in the cost of manipulation.

As an illustration of Theorem 4, consider the case where $n = 2$, and the sellers are identical, i.e., $A_1 = A_2$. Let $\iota(q) = q(2-q)\left(1 - \frac{2(1-q)^2}{2q^2-6q+3}\right)$. Let $q_{(1)}$ be the first real root of the equation $2q^3 - 10q^2 + 12q - 3 = 0$; $q_{(1)} \approx 0.339$. Equation (27) can be written as follows.

$$\underbrace{\frac{\iota(q_{(1)})}{b}}_{\approx \frac{0.152}{b}} < \lambda < \frac{q_i^{\mathsf{AM}}}{b} \implies \frac{d\pi_i^{\mathsf{PM}}}{d\lambda} > 0 \text{ for } i \in \{1,2\}.$$

Since q_i^{AM} is increasing in A_i, the above condition holds for high values of A_i. This result illustrates a paradoxical example akin to the *prisoner's dilemma*. Although sellers may be forced to manipulate in order to compete with others, every seller is better off had manipulation been preventable altogether. Since individual sellers do not benefit by unilaterally deviating from their equilibrium decision, all sellers are worse-off.

It is often argued that policing manipulation by sellers is an important activity of the platform. Often times, a platform can take measures to make manipulation more costly, and may hold the power to solve the dilemma. However, is it in the platform's interest to do so? If, empowered by platform technology, the consumers become sophisticated and are able to (partially) discern manipulation by sellers. How does this ability influence the outcome?

7.3 Effect of Consumer Sophistication and Platform Technology

In this section, we extend our main analysis to allow for consumer sophistication and the platform to be able to hinder manipulation. Suppose seller i manipulates by an amount

x_i. Instead of an increment of x_i in customer utility, suppose that their perceived quality increases by an amount δx_i, where $0 < \delta \leq 1$. That is, the consumer's (ex-ante) utility from seller i is:

$$u_i = a_i + \delta x_i - bp_i + \epsilon_i$$

δ can be interpreted in several ways. For example, δ corresponds to consumers' sophistication in detecting a seller's manipulation ($\delta = 1$ signifies complete naivety and $\delta = 0$ implies complete sophistication). δ can also be interpreted as a platform's technology that hinders a seller's manipulation. For example, several online platforms actively monitor the reviews posted for a seller and detect/flag potential fake reviews. The result below illustrates the effect of δ.

Theorem 5. *A seller's propensity to manipulate γ_i is increasing in consumer naivety δ. Consequently, for any δ, δ' such that $\delta \leq \delta'$, the set of sellers that manipulate under δ are contained in δ', i.e., $\mathcal{X}^{\delta} \subseteq \mathcal{X}^{\delta'}$. The market-share for no purchase option q_0^{PM} is decreasing in δ.*

While it can be verified that the market-share, markup and profits of sellers that do not manipulate decrease in consumer naivety, the same cannot be said about each seller that manipulates, due to a similar effect as discussed in Section 7.2. Intuitively, greater sophistication among consumers or a better platform technology is equivalent to an increase in the cost of manipulation. Consequently, as discussed in Section 7.2, an increase in consumer sophistication or a better platform technology to hinder manipulation may, paradoxically, benefit all sellers.

7.4 Implications for Platform's Revenue and Consumer Surplus

Suppose the platform uses a revenue-sharing/commission contract with the sellers. We investigate whether the platform benefits from manipulation. To do so, we compare industry revenues (i.e., the sum of revenues of all sellers) in the absence and presence of manipulation.

$$\begin{aligned} \text{Platform's Revenue} &= \text{Revenue Sharing Rate} \times \text{Industry Revenues}, \\ \text{where Industry Revenues} &\propto \sum_{i \in [n]} p_i q_i. \end{aligned}$$

Theorem 6. *Suppose $A_1 \leq A_2 \leq \ldots \leq A_n$ and sellers are homogeneous in their cost of manipulation. Further, suppose that the marginal production costs satisfy $c_1 \leq c_2 \leq \ldots \leq c_n$. Then, the industry revenues, and hence the platform's revenue, are higher in the presence of manipulation than that in its absence, i.e., $\sum_{i \in [n]} p_i^{\mathsf{PM}} q_i^{\mathsf{PM}} > \sum_{i \in [n]} p_i^{\mathsf{AM}} q_i^{\mathsf{AM}}$.*

Therefore, judging only from a single-shot revenue standpoint, the platform may have an interest in promoting manipulation, particularly if the nature of manipulation is innocuous, for example, by enticing good reviews via additional after-sale services and care. On the other hand, when manipulation involves untruthful, unethical, and even illegal means, the platform then has to evaluate the consequence of potential backlash and adverse effect on its platform users and its own reputation. While these considerations may also be relevant for individual sellers, they are more critical for the platform which depends more on return customers than one-time purchases. Toward that end, we examine the effect of manipulation on customer satisfaction, measured by the expected consumer surplus, in addition to revenue. It can be shown that the expected consumer surplus in the absence of manipulation is:

$$CS^{\mathsf{AM}} \quad = \quad \left(\gamma_e - \log q_0^{\mathsf{AM}}\right)/b$$

where γ_e is the Euler constant. A detailed derivation of consumer surplus under the MNL model under AM and PM is provided in Appendix D. In the presence of manipulation, we assume that the manipulation x_i influences the perceived quality but does not add to true consumer surplus. We derive the expected consumer surplus based on the *true* quality of the products:

$$CS^{\mathsf{PM}} \quad = \quad \left(\gamma_e - \log q_0^{\mathsf{PM}} - \overline{x}^{\mathsf{PM}}\right)/b.$$

where $\overline{x}^{\mathsf{PM}} \triangleq \sum_{i\in[n]} x_i^{\mathsf{PM}} q_i^{\mathsf{PM}}$ is proportional to the average level of manipulation in the market. Although it seems natural to predict that consumer surplus is negatively affected when manipulation is present, it is not always the case in the equilibrium.

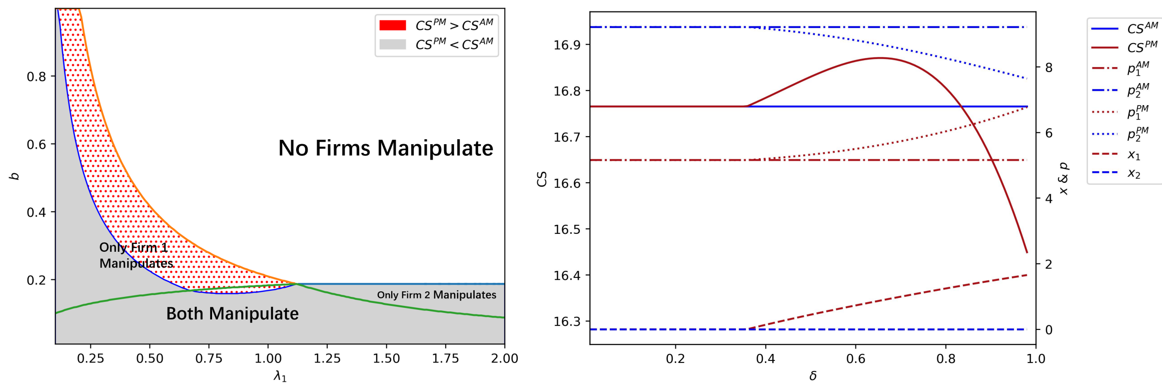
From the expressions of consumer surplus,

$$CS^{\mathsf{AM}} < CS^{\mathsf{PM}} \Leftrightarrow \overline{x}^{\mathsf{PM}} < \log\left(q_0^{\mathsf{AM}}/q_0^{\mathsf{PM}}\right).$$

Therefore, it is possible that allowing manipulation may either increase or decrease the expected consumer surplus, and determining the direction of the effect requires solving the respective equilibria and checking the condition $\overline{x}^{\mathsf{PM}} > \log\left(q_0^{\mathsf{AM}}/q_0^{\mathsf{PM}}\right)$.

In Figure 5, we present the outcomes in a duopoly with asymmetric firms that face heterogeneous costs of manipulation $h_i(x) = \lambda_i(e^x - 1)$, and the comparison of equilibrium consumer surplus under AM and PM. We vary the price-sensitivity parameter b and the consumer naivety parameter δ, plot the regions with different market outcomes and denote the region in which $CS^{\mathsf{PM}} > CS^{\mathsf{AM}}$ occurs (red dotted region) in Figure 5(a). Contrary to intuition, under certain circumstances, the presence of manipulation may improve consumer surplus.

Recall that sellers' response to competition is two-pronged: they may manipulate their perceived quality and/or adjust their price, depending on the cost of manipulation and consumer price sensitivity. In the example depicted, the two sellers have a notable difference in quality ($a_2 >> a_1$), so a dominant share of the consumer surplus is derived from seller 2. When price sensitivity is very low, both sellers tend to compete more via manipulation than via prices,



(a) $\delta = 1$

(b) $\lambda_1 = 0.3, b = 0.25$

Figure 5. Comparison of consumer surplus in an asymmetric duopoly under AM and PM with a heterogeneous cost of manipulation $h_i(x) = \lambda_i(e^x - 1)$. Values of Parameters: $a_1 = 1.2, a_2 = 3.2, c_1 = c_2 = 0.1, \lambda_2 = 3$.

and a greater extent of manipulation leads to higher prices by both sellers. Consequently, consumers are hurt (as shown in the gray region). On the other hand, when consumers are sufficiently price sensitive and λ_1 is small, while it is more efficient for seller 1 to compete via manipulation than price because of its low manipulation cost, seller 2 finds it expensive to manipulate, so they respond by choosing a lower price. Since the consumer surplus is predominantly derived from seller 2, a lower price by seller 2 contributes to an overall enhancement in total consumer welfare. In essence, the manipulation by the low-quality-low-manipulation-cost firm (seller 1) drives the high-quality-high-manipulation-cost firm (seller 2) to depress its price, leading to an increase in consumer surplus. Figure 5(b) focuses on this scenario and illustrates the compounding effect of δ (consumer naivety). Note that, in this specific instance, seller 2 refrains from engaging in manipulation due to a high cost of manipulation. Greater naivety of consumers allows for a higher extent of manipulation by seller 1, and leads to lower prices by seller 2. This drives up consumer surplus until seller 2's contribution to consumer surplus becomes less dominant, in which case we observe a trend reversal.

The examples above illustrate the mechanism for a possible positive effect of manipulation on consumer surplus – that is, manipulation by some sellers can create pricing pressure on other sellers and benefit consumers. We emphasize that, this is not to say that manipulation is beneficial to the market when consumers are not directly hurt; rather, sellers who do not manipulate are the ones who are unfairly placed at a disadvantage and forced to foot all damages of manipulation by others.

8 Conclusion

As consumers place increasing emphasis on online product reviews in purchasing decisions, sellers face strong pressure to elevate the rating of their product in order to compete with others. Consequently, sellers may be incentivized to adopt various means of review manipulations. Contradicting views exist in the literature regarding the sellers' tendency of review manipulation vis-à-vis the strength/quality/type of sellers. One view argues that high quality sellers have more to gain from manipulation and are more likely to manipulate whereas others present empirical evidence that sellers with low ratings exhibit stronger tendency to manipulate. We construct a model of multi-seller competition in which each seller sets its own price and review manipulation level to maximize profit. We solve for the unique equilibrium solution, and present a comprehensive characterization of the set of sellers that manipulate in equilibrium. We make several unique contributions to this literature: (i) by identifying an index γ directly computable from model parameters to scale each seller's relative propensity to manipulate, and proving that the set of sellers who manipulate is upward closed with respect to this propensity index, (ii) by partitioning the volume-rating space into regions that exhibit distinctive patterns of the iso-γ curve, equivalently, patterns of how manipulation propensity is affected by sellers' true quality, and (iii) by mapping our model of review manipulation to a star-rating system and illustrating how to apply it to a real-world data set. A key takeaway is that the two contradicting views regarding the relationship between seller quality and tendency to manipulate can be reconciled through our model and results: We establish the separation of the benefit-dominant region and cost-dominant region. In the cost-dominant region, low-quality sellers tend to manipulate, while in the benefit-dominant region, high-quality sellers tend to manipulate. Hence observations of which types of sellers manipulate in a given application or market may only reflect a censored snapshot view of a market. Decision makers need to be cautious in making generalizations.

Finally, we remark that the model we presented is potentially applicable to settings with *real* quality-enhancing efforts, such as implementing pre- or after-sale customer care and providing other value-adding perks. In this case, the equilibrium analysis and solution methods in Sections 3-5 carry through without technical difference and our model and methods serve as a novel contribution to this literature as well: The propensity-to-manipulate index, reinterpreted as propensity-to-enhance index, can similarly characterize the types of sellers who choose to engage in quality-enhancing efforts in an oligopoly with both price and quality competition and the effect of sellers equilibrium behavior. The discussion of consumer surplus in Section 7.2 certainly is specific to manipulating perceived quality instead of *real* quality, the latter of which is mathematically simpler to capture. The model in Section 6 applies specifically to the context of star-rating review manipulation and price competition and serves as a unique contribution to the literature.

References

Aksoy-Pierson, Margaret, Gad Allon, and Awi Federgruen, "Price competition under mixed multinomial logit demand functions," *Management Science*, 2013, *59* (8), 1817–1835.

Allon, Gad, Awi Federgruen, and Margaret Pierson, "Price Competition under Multinomial Logit Demand Functions with Random Coefficients," 2011.

Amazon Seller Central, "Amazon Services: Seller Forums," https://sellercentral.amazon.com/forums/t/incentivized-reviews/506637 2019. Accessed: August 4th, 2022.

Ananthakrishnan, Uttara M, Beibei Li, and Michael D Smith, "A tangled web: Should online review portals display fraudulent reviews?," *Information Systems Research*, 2020, *31* (3), 950–971.

Anderson, Simon P and André De Palma, "Multiproduct firms: A nested logit approach," *The Journal of Industrial Economics*, 1992, pp. 261–276.

Bernstein, Fernando and Awi Federgruen, "A general equilibrium model for industries with price and service competition," *Operations research*, 2004, *52* (6), 868–886.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 1995, pp. 841–890.

— , — , and — , "Differentiated products demand systems from a combination of micro and macro data: The new car market," *Journal of political Economy*, 2004, *112* (1), 68–105.

Berry, Steven T, "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 1994, pp. 242–262.

Besanko, David, Jean-Pierre Dubé, and Sachin Gupta, "Competitive price discrimination strategies in a vertical channel using aggregate retail data," *Management Science*, 2003, *49* (9), 1121–1138.

— , Sachin Gupta, and Dipak Jain, "Logit demand estimation under competitive pricing behavior: An equilibrium framework," *Management Science*, 1998, *44* (11-part-1), 1533–1547.

Chen, Le and Christo Wilson, "Observing algorithmic marketplaces in-the-wild," *ACM SIGecom Exchanges*, 2017, *15* (2), 34–39.

Chevalier, Judith A and Dina Mayzlin, "The effect of word of mouth on sales: Online book reviews," *Journal of Marketing Research*, 2006, *43* (3), 345–354.

Chevalier, Judith and Austan Goolsbee, "Measuring prices and price competition online: Amazon. com and BarnesandNoble. com," *Quantitative Marketing and Economics*, 2003, *1* (2), 203–222.

Chintagunta, Pradeep K, Shyam Gopinath, and Sriram Venkataraman, "The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets," *Marketing Science*, 2010, *29* (5), 944–957.

Crockett, Zachary, "5-star Phonies: Inside the Fake Amazon Review Complex," https://thehustle.co/amazon-fake-reviews/ 2019. Accessed: August 4th, 2022.

Dellarocas, Chrysanthos, "The digitization of word of mouth: Promise and challenges of online feedback mechanisms," *Management Science*, 2003, *49* (10), 1407–1424.

— , "Strategic manipulation of internet opinion forums: Implications for consumers and firms," *Management Science*, 2006, *52* (10), 1577–1593.

Fang, Youli, Hong Wang, Lili Zhao, Fengping Yu, and Caiyu Wang, "Dynamic knowledge graph based fake-review detection," *Applied Intelligence*, 2020, *50*, 4281–4295.

Farahat, Amr and Georgia Perakis, "A comparison of Bertrand and Cournot profits in oligopolies with differentiated products," *Operations Research*, 2011, *59* (2), 507–513.

Federal Trade Commission, "Soliciting and Paying for Online Reviews: A Guide for Marketers," 2022.

— , "Federal Trade Commission Announces Proposed Rule Banning Fake Reviews and Testimonials," 2023.

Gallego, Guillermo and Ruxian Wang, "Multiproduct price optimization and competition under the nested logit model with product-differentiated price sensitivities," *Operations Research*, 2014, *62* (2), 450–461.

— **, Woonghee Tim Huh, Wanmo Kang, and Robert Phillips**, "Price competition with the attraction demand model: Existence of unique equilibrium and its stability," *Manufacturing & Service Operations Management*, 2006, *8* (4), 359–375.

He, Sherry, Brett Hollenbeck, and Davide Proserpio, "The market for fake reviews," *Marketing Science*, 2022.

Lee, Chung-Seung and Metin Çakanyildirim, "Price Competition Under Mixed Multinomial Logit Demand: Sufficiency Conditions for Validating the Model," *Production and Operations Management*, 2021, *30* (9), 3272–3283.

Li, Hongmin and Woonghee Tim Huh, "Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications," *Manufacturing & Service Operations Management*, 2011, *13* (4), 549–563.

Luca, Michael, "Reviews, Reputation, and Revenue: The Case of Yelp. com," 2011.

— **and Georgios Zervas**, "Fake it till you make it: Reputation, competition, and Yelp review fraud," *Management Science*, 2016, *62* (12), 3412–3427.

Mayzlin, Dina, "Promotional chat on the Internet," *Marketing Science*, 2006, *25* (2), 155–163.

— **, Yaniv Dover, and Judith Chevalier**, "Promotional reviews: An empirical investigation of online review manipulation," *American Economic Review*, 2014, *104* (8), 2421–55.

McFadden, Daniel et al., "Conditional logit analysis of qualitative choice behavior," 1973.

McLachlan, Geoffrey J and Thriyambakam Krishnan, *The EM algorithm and extensions*, John Wiley & Sons, 2007.

Mohawesh, Rami, Shuxiang Xu, Son N Tran, Robert Ollington, Matthew Springer, Yaser Jararweh, and Sumbal Maqsood, "Fake reviews detection: A survey," *IEEE Access*, 2021, *9*, 65771–65802.

Salminen, Joni, Chandrashekhar Kandpal, Ahmed Mohamed Kamel, Soon gyo Jung, and Bernard J Jansen, "Creating and detecting fake reviews of online products," *Journal of Retailing and Consumer Services*, 2022, *64*, 102771.

Sun, Monic, "How does the variance of product ratings matter?," *Management Science*, 2012, *58* (4), 696–707.

Techcrunch, "Amazon bans incentivized reviews tied to free or discounted products," 2017.

The Wall Street Journal, "You're Probably Falling for Fake Product Reviews," 2023.

Vulcano, Gustavo, Garrett Van Ryzin, and Richard Ratliff, "Estimating primary demand for substitutable products from sales transaction data," *Operations Research*, 2012, *60* (2), 313–334.

Wang, Ruxian, Chenxu Ke, and Shiliang Cui, "Product price, quality, and service decisions under consumer choice models," *Manufacturing & Service Operations Management*, 2022, *24* (1), 430–447.

Wang, Xin, Feng Mai, and Roger HL Chiang, "Database submission—market dynamics and user-generated content about tablet computers," *Marketing Science*, 2014, *33* (3), 449–458.

World Economic Forum, "Fake online reviews cost 152 billion dollars a year. Here's how e-commerce sites can stop them," 2021.

Yelp, "Why Yelp Doesn't Condone Review Solicitation," 2017.

A Proofs of Technical Results

This appendix provides the proofs of all technical results in the main paper. Besides this appendix, we also provide a supplementary appendix, Appendix B, that provides some helpful results.

A.1 Proof of Results in Section 4

Proof of Theorem 1. For completeness, we show firm i's best-response. We then solve for the equilibrium outcome.

The derivative of π_i w.r.t m_i is:

$$\frac{d\pi_i}{dm_i} = q_i + m_i\left(\frac{dq_i}{dm_i}\right)$$

From (5), we have

$$\frac{dq_i}{dm_i} = \frac{A_i e^{-bm_i}(-b)\left(1+\sum_{i\in[n]} A_i e^{-bm_i}\right) - A_i e^{-bm_i} A_i e^{-bm_i}(-b)}{\left(1+\sum_{i\in[n]} A_i e^{-bm_i}\right)^2} = -bq_i(1-q_i).$$

Substituting this back in the RHS of $\frac{d\pi_i}{dm_i}$:

$$\frac{d\pi_i}{dm_i} = q_i\left(1 - bm_i(1-q_i)\right).$$

Consider the term inside the brackets in the RHS above. This term is decreasing in m_i. To see this,

$$bm_i(1-q_i) = bm_i\left(\frac{1+\sum_{j\neq i} A_j e^{-bm_j}}{1+\sum_{j\neq i} A_j e^{-bm_j} + A_i e^{-bm_i}}\right)$$

Both terms in the RHS above are increasing in m_i; therefore, $1-bm_i(1-q_i)$ is decreasing in m_i. For given $\mathbf{m}_{-i}$, let $m_i(\mathbf{m}_{-i})$ denote the unique value of m_i s.t. $1-bm_i(1-q_i)=0$. Then, for fixed $\mathbf{m}_{-i}$, $\frac{d\pi_i}{dm_i} > 0$ iff $m_i < m_i(\mathbf{m}_{-i})$. That is, π_i is unimodal in m_i; at optimality, $m_i = m_i(\mathbf{m}_{-i}) = \frac{1}{b(1-q_i)}$.

To solve for $m_i(\mathbf{m}_{-i})$, we can write the $m_i(\mathbf{m}_{-i})$ as:

$$\begin{aligned}
bm_i &= \frac{1}{1-q_i} \implies bm_i - 1 = \frac{q_i}{\underbrace{1-q_i}_{q_0+\sum_{j\neq i} q_j}} = \frac{A_i e^{-bm_i}}{1+\sum_{j\neq i, j\in[n]} A_j e^{-bm_j}} \\
\implies m_i(\mathbf{m}_{-i}) &= \frac{1}{b}\left(1 + \frac{A_i e^{-bm_i}}{1+\sum_{j\neq i, j\in[n]} A_j e^{-bm_j}}\right). \qquad (28)
\end{aligned}$$

The RHS is strictly decreasing in m_i. Thus, a unique fixed point to the RHS exists. Further,

$$\begin{aligned}
(bm_i - 1)e^{bm_i - 1} &= \frac{A_i e^{-1}}{1+\sum_{j\neq i, j\in[n]} A_j e^{-bm_j}} \implies bm_i - 1 = \mathcal{W}\left(\frac{A_i e^{-1}}{1+\sum_{j\neq i, j\in[n]} A_j e^{-bm_j}}\right) \\
\implies m_i(\mathbf{m}_{-i}) &= \frac{1}{b}\left(1 + \mathcal{W}\left(\frac{A_i e^{-1}}{1+\sum_{j\neq i, j\in[n]} A_j e^{-bm_j}}\right)\right)
\end{aligned}$$

where $\mathcal{W}(\cdot)$ denotes the Lambert-W function.[9]

We rewrite (5) as follows:

$$q_i = A_i e^{-bm_i} q_0,$$

[9] Fix $k > 0$. The equation $ze^z = k$ has a unique solution at $z = \mathcal{W}(k)$. Further,

$$\frac{dz}{dk} = \left.\frac{1}{e^z(1+z)}\right|_{z=\mathcal{W}(k)} = \frac{1}{e^{\mathcal{W}(k)} + k}.$$

where $q_0 = \frac{1}{1+\sum_{j\in[n]} A_j e^{x_j - bm_j}}$. Substituting for m_i from (6) in the RHS above, we have:

$$q_i e^{\frac{1}{1-q_i}} = A_i q_0 \implies q_i = f^{-1}(A_i q_0)$$

Below, we identify the equilibrium q_0.

$$q_0 = 1 - \sum_{j\in[n]} q_j \implies q_0 = 1 - \sum_{j\in[n]} f^{-1}(A_i q_0)$$

Since $f(x)$ is increasing in x, with $f(0) = 0$ and $f(1) = \infty$, the RHS is decreasing in q_0; thus, the solution to q_0 exists and is unique. Combining the solution to q_0 and (6), we have the required result. □

A.2 Proofs of Results in Section 5

Proof of Lemma 1. Fix $\mathbf{x}$. Let $\breve{A}_j$ denote the following.

$$\breve{A}_j = A_j e^{x_j}.$$

Observe that the above problem is identical to that in Section 4, except for $A_j \to \breve{A}_j$. For fixed $\mathbf{m}_{-i}, \mathbf{x}$, the unimodality of π_i in m_i follows from (4) and the above observation. Therefore, FOC's identify the optimal m_i. Using (28), we have:

$$\frac{\partial \pi_i}{\partial m_i} = q_i + m_i \frac{\partial q_i}{\partial m_i} = 0 \quad \implies \quad bm_i - 1 = \frac{\breve{A}_i e^{-bm_i}}{1+\sum_{j\neq i} \breve{A}_j e^{-bm_j}} = \frac{A_i e^{x_i - bm_i}}{1+\sum_{j\neq i} A_j e^{x_j - bm_j}}$$

$$\implies \quad m_i = \frac{1}{b}\left(\underbrace{1 + \frac{A_i e^{x_i - bm_i}}{1+\sum_{j\neq i} A_j e^{x_j - bm_j}}}_{\frac{1}{1-q_i}}\right).$$

For fixed $(\mathbf{m}_{-i}, \mathbf{x})$, (10) has a unique solution for m_i. Furthermore, (10) can be written as:

$$(bm_i - 1)e^{bm_i - 1} = \frac{A_i e^{x_i - 1}}{\left(1+\sum_{j\neq i} A_j e^{x_j - bm_j}\right)} \implies m_i = \frac{1}{b}\left[1 + \mathcal{W}\left(\frac{A_i e^{x_i - 1}}{1+\sum_{j\neq i} A_j e^{x_j - bm_j}}\right)\right]. \tag{29}$$

From (29), m_i is increasing in x_i, increasing in $\mathbf{m}_{-i}$, and decreasing in $\mathbf{x}_{-i}$.

Further, $x_i - bm_i(x_i)$ is increasing in x_i. To see this, consider the derivative of $x_i - bm_i(x_i)$:

$$\frac{d}{dx_i}(x_i - bm_i) = -b\frac{dm_i}{dx_i},$$

We show that the RHS is positive. Consider the second term. Using (10),

$$\begin{aligned} b\frac{dm_i}{dx_i} &= \frac{A_i e^{x_i - bm_i}}{1+\sum_{j\neq i} A_j e^{x_j - bm_j}}\left(1 - b\frac{dm_i}{dx_i}\right) \\ \implies b\frac{dm_i}{dx_i} &= \frac{A_i e^{x_i - bm_i}}{1+\sum_{j\in[n]} A_j e^{x_j - bm_j}} = q_i. \end{aligned}$$

Since $(1 - q_i) > 0$ we have that $x_i - bm_i$ is increasing in x_i. □

Proof of Lemma 2. Since $\pi_i = m_i q_i - h(x_i)$, the marginal effect of manipulation on seller i's profits (i.e., derivative of π_i w.r.t x_i) is:

$$\frac{d\pi_i}{dx_i} = \underbrace{q_i \frac{dm_i}{dx_i}}_{\text{effect of manipulation on infra-marginal units}} + \underbrace{m_i \frac{dq_i}{dx_i}}_{\text{effect of manipulation on the marginal unit}} - \underbrace{h_i'(x_i)}_{\text{marginal cost of manipulation}}. \tag{30}$$

The first term in the RHS of (30) is the effect on the inframarginal units, while the second term is the effect on the marginal unit. First, using (10), we write m_i as follows:

$$(bm_i - 1)e^{bm_i - 1} = \frac{A_i e^{x_i - 1}}{1 + \sum_{j \neq i} A_j e^{x_j - bm_j}}.$$

Differentiating both sides w.r.t. x_i, we have:

$$\begin{aligned}
\frac{d}{dx_i}\left((bm_i - 1)e^{bm_i - 1}\right) &= \frac{d}{dx_i}\left(\frac{A_i e^{x_i - 1}}{1 + \sum_{j \neq i} A_j e^{x_j - bm_j}}\right) \\
\implies \underbrace{\left(b^2 m_i e^{bm_i - 1}\right)}_{= \frac{\partial}{\partial m_i}\left((bm_i - 1)e^{bm_i - 1}\right)} \frac{dm_i}{dx_i} &= \frac{A_i e^{x_i - 1}}{1 + \sum_{j \neq i} A_j e^{x_j - bm_j}} \\
\implies \frac{dm_i}{dx_i} &= \frac{1}{b^2 m_i}\left(\frac{A_i e^{x_i - bm_i}}{1 + \sum_{j \neq i} A_j e^{x_j - bm_j}}\right) \\
&= \frac{1}{b^2 m_i}\left(\frac{q_i}{1 - q_i}\right) = \frac{q_i}{b}.
\end{aligned}$$

Next, we can write $\frac{dq_i}{dx_i}$ as follows:

$$\frac{dq_i}{dx_i} = \frac{\partial q_i}{\partial x_i} + \frac{\partial q_i}{\partial m_i}\frac{dm_i}{dx_i}.$$

Using (3),

$$\begin{aligned}
\frac{\partial q_i}{\partial x_i} &= \frac{(A_i e^{x_i - bm_i})(1 + \sum_j A_j e^{x_j - bm_j}) - (A_i e^{x_i - bm_i})(A_i e^{x_i - bm_i})}{(1 + \sum_j A_j e^{x_j - bm_j})^2} \\
&= q_i(1 - q_i). \\
\frac{\partial q_i}{\partial m_i} &= \frac{(A_i e^{x_i - bm_i}(-b))(1 + \sum_j A_j e^{x_j - bm_j}) - (A_i e^{x_i - bm_i})(A_i e^{x_i - bm_i}(-b))}{(1 + \sum_j A_j e^{x_j - bm_j})^2} \\
&= -bq_i(1 - q_i).
\end{aligned}$$

Therefore,

$$\frac{dq_i}{dx_i} = q_i(1 - q_i)^2. \tag{31}$$

Combining the above, (30) simplifies to:

$$\frac{d\pi_i}{dx_i} = \left(\frac{q_i^2}{b} + m_i q_i (1 - q_i)^2\right) - h_i'(x_i)$$

Substituting for m_i from (10), we have:

$$\frac{d\pi_i}{dx_i} = \underbrace{\frac{q_i}{b}}_{\text{marginal benefit from manipulation}} - \underbrace{h_i'(x_i)}_{\text{marginal cost of manipulation}}. \tag{32}$$

The critical point(s), denoted by x_i^*, solve:

$$q_i(x_i, m_i(x_i)) = bh_i'(x_i). \tag{33}$$

At the critical point(s), $x_i = x_i^*$, the second derivative is:

$$\begin{aligned}
\left.\frac{d^2\pi}{dx_i^2}\right|_{x_i = x_i^*} &= \frac{d}{dx_i}\left(\frac{d\pi_i}{dx_i}\right)\Bigg|_{x_i = x_i^*} = \left(\frac{q_i(1 - q_i)^2}{b} - h_i''(x_i)\right)\Bigg|_{x_i = x_i^*} \\
&= h_i'(x_i^*)(1 - q_i)^2 - h_i''(x_i^*) < 0. \qquad \text{(due to Assumption 1(b))}
\end{aligned}$$

Therefore, $\pi_i(x_i)$ is quasi-concave. Consequently, the following hold:

1. If $\frac{q_i(0)}{b} < h_i'(0)$, then, $\frac{d\pi_i}{dx_i} < 0$ for all $x_i > 0$.
2. If $\frac{q_i(0)}{b} > h_i'(0)$, then, FOC's identify the unique interior maximum.

For convenience, define $\breve{f}(z)$, $z \in [0,1)$, as follows:

$$\breve{f}(z) \triangleq \left(\frac{z}{1-z}\right) e^{\frac{1}{1-z}}. \tag{34}$$

$\breve{f}(z)$ is increasing in z. The solution to the FOC in (33) can be expressed as follows.

$$\underbrace{\breve{f}^{-1}\left(\frac{A_i e^{x_i}}{1+\sum_{j\neq i} A_j e^{x_j - bm_j}}\right)}_{q_i(x_i)} = bh_i'(x_i). \tag{35}$$

While the LHS and RHS are both increasing in x_i, since π_i is quasi-concave in x_i, it follows that the above equation has a unique solution for x_i. □

Proof of Theorem 2. From Lemmas 1 and 2, we have seller i's choice of m_i and x_i. Depending on x_i^{PM}, one of the following applies to seller i:

- If $x_i^{\mathsf{PM}} = 0$ (i.e., $i \in \mathcal{X}^{\mathsf{C}}$), then, using (10), we have the following:

$$\begin{aligned} q_i &= \frac{A_i e^{-\frac{1}{1-q_i}}}{1+\sum_j A_j e^{x_j^{\mathsf{PM}} - bm_j^{\mathsf{PM}}}} = A_i e^{-\frac{1}{1-q_i}} q_0 \\ \implies \underbrace{q_i e^{\frac{1}{1-q_i}}}_{f(q_i)} &= A_i q_0 \implies q_i = f^{-1}(A_i q_0). \end{aligned}$$

- If $x_i^{\mathsf{PM}} > 0$ (i.e., $i \in \mathcal{X}$), then, using (10) and (11), we have the following:

$$\begin{aligned} q_i &= \frac{A_i e^{h_i'^{-1}\left(\frac{q_i}{b}\right) - \frac{1}{1-q_i}}}{1+\sum_j A_j e^{x_j^{\mathsf{PM}} - bm_j^{\mathsf{PM}}}} = A_i e^{h_i'^{-1}\left(\frac{q_i}{b}\right) - \frac{1}{1-q_i}} q_0 \\ \implies \underbrace{q_i e^{\frac{1}{1-q_i} - h_i'^{-1}\left(\frac{q_i}{b}\right)}}_{g_i(q_i)} &= A_i q_0 \implies q_i = g_i^{-1}(A_i q_0). \end{aligned}$$

We solve for the equilibrium value of q_0. Since $q_0 = 1 - \sum_{i\in[n]} q_i$, we have:

$$q_0 = 1 - \sum_{i\in\mathcal{X}} g_i^{-1}(A_i q_0) - \sum_{i\in\mathcal{X}^{\mathsf{C}}} f^{-1}(A_i q_0).$$

Since $f(\cdot)$ and $g_i(\cdot)$, $i \in [n]$ are increasing functions, and $f(0) = g(0) = 0$ and $f(1) = g(1) = \infty$, the RHS is decreasing in q_0. Further, at $q_0 = 0$ (resp., $q_0 = 1$), the LHS is strictly smaller (resp., larger) than the RHS. Hence, the above equation has a unique solution for $q_0 \in (0,1)$. Denote this solution by q_0^{PM}. Substituting q_0^{PM} in the expression for q_i, we have

$$q_i^{\mathsf{PM}} = \begin{cases} f^{-1}(A_i q_0^{\mathsf{PM}}), & \text{if } i \in \mathcal{X}^{\mathsf{C}}; \\ g_i^{-1}(A_i q_0^{\mathsf{PM}}), & \text{if } i \in \mathcal{X}. \end{cases} \tag{36}$$

Substituting q_i^{PM} in (10) and (11), we have:

$$m_i = \frac{1}{b(1-q_i^{\mathsf{PM}})} \text{ and } x_i^{\mathsf{PM}} = \begin{cases} 0, & \text{if } i \in \mathcal{X}^{\mathsf{C}}; \\ h_i'^{-1}\left(\frac{q_i^{\mathsf{PM}}}{b}\right), & \text{if } i \in \mathcal{X}. \end{cases}$$

□

Proof of Lemma 1. Consider the equilibrium outcome under PM (the equilibrium quantities are denoted by the superscript PM). Recall the definition of $\mathcal{X}$ in (12). From Assumption 2, we have that, in equilibrium,

$\exists i \in [n]$ s.t. $x_i^{\mathsf{PM}} > 0$. That is, $\mathcal{X} \neq \emptyset$. Also, recall the definition of γ_i in (15), and the equilibrium outcome (the market share, profit margin and the extent of manipulation by each firm) under PM in Theorem 2.

First, recall from (13) that if $h_i'(0) > \frac{1}{b}$, then it is a dominant strategy for firm i to not manipulate. From (15), we have that $\gamma_i = 0$ if $h_i'(0) > \frac{1}{b}$. Combining these two statements, we have $\gamma_i = 0 \implies i \in \mathcal{X}^{\mathsf{C}}$. To prove our result, it then suffices to restrict attention to the set of of firms s.t their γ is strictly positive.

Consider two such firms, say i and j s.t. $0 < \gamma_j \leq \gamma_i$. We show the following two claims.

(a) Suppose $i \in \mathcal{X}^{\mathsf{C}}$. Then, $j \in \mathcal{X}^{\mathsf{C}}$.

(b) Suppose $j \in \mathcal{X}$. Then, $i \in \mathcal{X}$.

Consider part (a). Since firm $i \in \mathcal{X}^{\mathsf{C}}$ (i.e., $x_i^{\mathsf{PM}} = 0$), the following holds:

$$\begin{aligned} i \in \mathcal{X}^{\mathsf{C}} \quad &\implies \quad h_i'(0) \geq \frac{q_i^{\mathsf{PM}}}{b} \\ &\implies \quad bh_i'(0) \geq f^{-1}(A_i q_0^{\mathsf{PM}}) \\ &\implies \quad f\left(h_i'(0)\right) \geq A_i q_0^{\mathsf{PM}} \\ &\implies \quad \frac{1}{q_0^{\mathsf{PM}}} \geq \gamma_i. \end{aligned}$$

Since $\gamma_j \leq \gamma_i$, it follows that $\gamma_j \leq \frac{1}{q_0^{\mathsf{PM}}}$, i.e.,

$$f\left(bh_j'(0)\right) \geq A_j q_0^{\mathsf{PM}}$$

We will show that $j \in \mathcal{X}^{\mathsf{C}}$ by contradiction. Suppose $j \in \mathcal{X}$. Then, it must be the case that $h_j'(0) < \frac{1}{b} q_j^{\mathsf{PM}}$ (i.e., $x_j > 0$). Recall from Theorem 2 that if $j \in \mathcal{X}$, then $q_j^{\mathsf{PM}} = g_j^{-1}(A_j q_0^{\mathsf{PM}})$. Using the definition of $g_j(\cdot)$ in (14) and the fact that $g_j(\cdot)$ is monotone, we have:

$$\begin{aligned} bh_j'(0) < q_j^{\mathsf{PM}} \quad &\implies \quad g_j\left(bh_j'(0)\right) < g_j(q_j^{\mathsf{PM}}) \text{ (since } g_j(\cdot) \text{ is monotone)}, \\ &\implies \quad f\left(bh_j'(0)\right) < A_j q_0^{\mathsf{PM}} \text{ (since } g_j\left(bh_j'(0)\right) = f\left(bh_j'(0)\right) \text{ and } q_j^{\mathsf{PM}} = g_j^{-1}(A_j q_0^{\mathsf{PM}})), \\ &\implies \quad \frac{1}{q_0^{\mathsf{PM}}} < \gamma_j. \end{aligned}$$

which is a contradiction. Thus, $j \in \mathcal{X}^{\mathsf{C}}$.

Consider part (b). Since $j \in \mathcal{X}$ (i.e., $x_j^{\mathsf{PM}} > 0$), the following holds:

$$\begin{aligned} j \in \mathcal{X} \quad &\implies \quad bh_j'(0) < q_j^{\mathsf{PM}} \\ &\implies \quad bh_j'(0) < g_j^{-1}(A_j q_0^{\mathsf{PM}}) \\ &\implies \quad g_j\left(bh_j'(0)\right) < A_j q_0^{\mathsf{PM}} \\ &\implies \quad f\left(bh_j'(0)\right) < A_j q_0^{\mathsf{PM}} \quad \text{(since } g_j\left(bh_j'(0)\right) = f\left(bh_j'(0)\right)), \\ &\implies \quad \frac{1}{q_0^{\mathsf{PM}}} < \gamma_j. \end{aligned}$$

Since $\gamma_i \geq \gamma_j$, it follows that $\gamma_i > \frac{1}{q_0^{\mathsf{PM}}}$, i.e., $f\left(bh_i'(0)\right) < A_i q_0^{\mathsf{PM}}$. We will show that $i \in \mathcal{X}$ by contradiction. Suppose $i \in \mathcal{X}^{\mathsf{C}}$. Then, it must be the case that $h_i'(0) \geq \frac{q_i^{\mathsf{PM}}}{b}$. From Theorem 2, since $i \in \mathcal{X}^{\mathsf{C}}$, $q_i^{\mathsf{PM}} = f^{-1}(A_i q_0^{\mathsf{PM}})$. Substituting the above, we have $f(bh_i'(0)) \geq A_i q_0^{\mathsf{PM}}$, or equivalently, $\frac{1}{q_0^{\mathsf{PM}}} \geq \gamma_i$, which is a contradiction. Therefore, $i \in \mathcal{X}$. □

Proof of Lemma 3. Consider any $i \in [n-1]$. Below, we prove (a), i.e., $x_i^{\mathsf{PM}} \leq x_{i+1}^{\mathsf{PM}}$. Since q_i is increasing in x_i, part (b) follows. Since m_i^{PM} is increasing in q_i, part (c) follows.

We show (a) by contradiction. Suppose that $x_{i+1}^{\mathsf{PM}} < x_i^{\mathsf{PM}}$. We divide the analysis into two cases:

- Suppose $0 < x_{i+1}^{\mathsf{PM}} < x_i^{\mathsf{PM}}$. Since the cost of manipulation is homogeneous across sellers,

$$\begin{aligned} x_{i+1}^{\mathsf{PM}} < x_i^{\mathsf{PM}} \quad &\implies \quad h'^{-1}\left(\frac{q_{i+1}^{\mathsf{PM}}}{b}\right) < h'^{-1}\left(\frac{q_i^{\mathsf{PM}}}{b}\right) \\ &\implies \quad q_{i+1}^{\mathsf{PM}} < q_i^{\mathsf{PM}} \\ &\implies \quad g^{-1}(A_{i+1}q_0^{\mathsf{PM}}) < g^{-1}(A_i q_0^{\mathsf{PM}}) \text{ (from (36))} \\ &\implies \quad A_{i+1} < A_i, \text{which is a contradiction.} \end{aligned}$$

- Suppose $0 = x_{i+1}^{\mathsf{PM}} < x_i^{\mathsf{PM}}$. It must be the case that

$$\frac{q_{i+1}^{\mathsf{PM}}}{b} \leq h'(0) < \frac{q_i^{\mathsf{PM}}}{b}.$$

  Rewriting the above,

$$\begin{aligned} q_{i+1}^{\mathsf{PM}} \leq bh'(0) < q_i^{\mathsf{PM}} \quad &\implies \quad f^{-1}(A_{i+1}q_0^{\mathsf{PM}}) \leq bh'(0) < g^{-1}(A_i q_0^{\mathsf{PM}}) \\ &\implies \quad A_{i+1}q_0^{\mathsf{PM}} \leq f\left(bh'(0)\right) = g\left(bh'(0)\right) < A_i q_0^{\mathsf{PM}} \\ &\implies \quad A_{i+1} < A_i, \text{ which is a contradiction.} \end{aligned}$$

We now compare seller profits, i.e., we show $\pi_i^{\mathsf{PM}} \leq \pi_{i+1}^{\mathsf{PM}}$.

- Suppose $x_i^{\mathsf{PM}} > 0$ (equivalently, $\frac{q_i^{\mathsf{PM}}}{b} > h'(0)$). Then, seller i's profit can be written as:

$$\pi_i = m_i q_i - h(x_i) = \frac{q_i}{b(1-q_i)} - h\left(h'^{-1}\left(\frac{q_i}{b}\right)\right). \tag{37}$$

  The RHS is increasing in q_i. To see this,

$$\frac{d\pi_i}{dq_i} = \frac{1}{b}\left(\frac{1}{(1-q_i)^2} - \frac{\frac{q_i}{b}}{h''(h'^{-1}(\frac{q_i}{b}))}\right)$$

  The first term inside the brackets is strictly larger than 1, while the second term inside the brackets is smaller than 1 (from Assumption 1(b)). To see this, let $\hat{y} = h'^{-1}(\frac{q_i}{b})$. The term inside the bracket can be written as $\frac{1}{(1-q_i)^2} - \frac{h'(\hat{y})}{h''(\hat{y})}$. Therefore, the RHS is positive. Therefore, $q_i^{\mathsf{PM}} \leq q_{i+1}^{\mathsf{PM}} \implies \pi_i^{\mathsf{PM}} \leq \pi_{i+1}^{\mathsf{PM}}$.

- Suppose $x_i^{\mathsf{PM}} = 0$ (equivalently, $\frac{q_i^{\mathsf{PM}}}{b} \leq h'(0)$). Then, seller i's profit can be written as:

$$\pi_i = m_i q_i = \frac{q_i}{b(1-q_i)} \tag{38}$$

  The RHS is increasing in q_i. We have the following two cases, depending on the value of x_{i+1}^{PM}:

  - Suppose $x_{i+1}^{\mathsf{PM}} = 0$. Since π_i is increasing in q_i, and $q_i^{\mathsf{PM}} \leq q_{i+1}^{\mathsf{PM}}$, we have that $\pi_i^{\mathsf{PM}} \leq \pi_{i+1}^{\mathsf{PM}}$.
  - Suppose $x_{i+1}^{\mathsf{PM}} > 0 (= x_i^{\mathsf{PM}})$. Observe that:

$$\pi_i^{\mathsf{PM}} = \frac{q_i^{\mathsf{PM}}}{b(1-q_i^{\mathsf{PM}})} < \left(\frac{q}{b(1-q)}\right)\bigg|_{q=bh'(0)} < \frac{q_{i+1}^{\mathsf{PM}}}{b(1-q_{i+1}^{\mathsf{PM}})} - h\left(h'^{-1}\left(\frac{q_{i+1}^{\mathsf{PM}}}{b}\right)\right) = \pi_{i+1}^{\mathsf{PM}}.$$

    The first inequality follows from (38) and the second inequality follows from (37).

  In other words, the equilibrium profit π_i is increasing in q_i in both cases (see (37) and (38)), and is continuous at the point of non-differentiability (i.e., $q_i = bh'(0)$).

$\square$

Proof of Lemma 4. Recall the definition of γ_i:

$$\gamma_i = \frac{A_i}{f\left(bh_i'(0)\right)}$$

Since $A_i = A$ for all $i \in [n]$ and $h_1'(x) \geq h_2'(x) \geq \ldots h_n'(x)$, we have that $\gamma_1 \leq \gamma_2 \leq \ldots \gamma_n$. Since $\mathcal{X}$ is upward-closed in γ_i, we have that $\exists\ i^*$ s.t. $\mathcal{X}^{\mathsf{C}} = \{1, 2, \ldots, i^* - 1\}$ and $\mathcal{X} = \{i^*, i^* + 1, n\}$.

Consider $i \in \mathcal{X}^{\mathsf{C}}$. We have that $q_i^{\mathsf{PM}} = f^{-1}(Aq_0^{\mathsf{PM}})$. Hence, $q_i = q_j$ for any $i, j \in \mathcal{X}^{\mathsf{C}}$.

Consider $i \in \mathcal{X}$. We have that $g_i(q) = f(q)e^{-h_i'^{-1}\left(\frac{q}{b}\right)}$. Since $h_1'(x) \geq h_2'(x) \ldots h_n'(x)$, it follows that $g_i(x) \geq g_{i+1}(x)$. Therefore, $q_i^{\mathsf{PM}} \leq q_{i+1}^{\mathsf{PM}}$ for $i \in \mathcal{X}$.

Since m_i is monotone in q_i, the comparison of m_i follows. For $i \in \mathcal{X}$, since $x_i = h_i'^{-1}\left(\frac{q_i}{b}\right)$, it follows that $x_i^{\mathsf{PM}} \leq x_{i+1}^{\mathsf{PM}}$. □

Proof of Lemma 5. Substituting for $h_i(x, A) = \mathcal{H}(A)h(x)$ in the definition of γ_i from (15), we have:

$$\gamma_i = \frac{A_i}{f\left(b\mathcal{H}(A_i)h'(0)\right)}.$$

Therefore,

$$\frac{d\gamma_i}{dA_i} = \frac{A_i}{f(z)}\left(\frac{1}{A_i} - \frac{f'(z)}{f(z)}bh'(0)\mathcal{H}'(A_i)\right)$$

where $z = bh'(0)\mathcal{H}(A_i)$. Further, $0 < z < 1$. In the RHS, the term outside the bracket is positive. It suffices to focus on the term within the brackets. The term inside the brackets is positive iff the following holds:

$$\begin{aligned} \left(\frac{1}{z} + \frac{1}{(1-z)^2}\right) z\frac{\mathcal{H}'(A_i)}{\mathcal{H}(A_i)} < \frac{1}{A_i} \quad &\Leftrightarrow \quad \frac{\frac{\mathcal{H}'(A_i)}{\mathcal{H}(A_i)}}{\frac{1}{A_i}} < \frac{f(z)}{zf'(z)} \\ &\Leftrightarrow \quad \underbrace{\frac{\partial \log \mathcal{H}(A_i)}{\partial \log A_i}}_{\varepsilon_A} < \frac{1}{1 + \frac{z}{(1-z)^2}} \end{aligned}$$

Since the RHS is strictly less than 1, it holds that if $\varepsilon_A > 1$, then, γ_i is decreasing in A_i. □

A.3 Proofs of Results in Section 6

Proof of Lemma 6. Observe from equation (21) that the quantity $bh_i'(0)$ can be written as

$$bh_i'(0) = \frac{k_1 v_i^{\mathsf{tr}}(-\beta_p)}{(R - r_i^{\mathsf{tr}})\beta_r} = \frac{v_i^{\mathsf{tr}}}{\overline{\overline{v}} - \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1}}.$$

where $\overline{\overline{v}} = \frac{\beta_r R}{(-\beta_p)k_1}$ is as shown in (22). First, recall from (15) that if $bh_i'(0) > 1$, then $\gamma_i = 0$. The condition can be equivalently stated as $v_i^{\mathsf{tr}} > \overline{\overline{v}} - \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1}$. Rearranging the terms, we have:

$$\gamma_i = 0 \Leftrightarrow v_i^{\mathsf{tr}} + \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1} > \overline{\overline{v}}.$$

Therefore, if $v_i^{\mathsf{tr}} > \overline{\overline{v}}$, then, $\gamma_i = 0$ for all r_i^{tr}.

If the above condition does not hold, then γ_i can be written as:

$$\gamma_i = \frac{A_i}{f(bh_i'(0))} = \frac{\exp(\beta_0 + \beta_r r_i^{\mathsf{tr}} + \beta_p c_i)}{f\left(\frac{v_i^{\mathsf{tr}}}{\overline{\overline{v}} - \frac{r_i^{\mathsf{tr}}\beta_r}{(-\beta_p)k_1}}\right)}$$

Since the RHS is decreasing in v_i^{tr} for $v_i^{\mathsf{tr}} < \overline{\overline{v}} - \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1}$, we have part (a) of the result.

Next, we show part (b), i.e., we understand how γ_i changes with r_i^{tr}. For convenience, let $\star$ denote the term $\left(\frac{v_i^{\mathsf{tr}}}{\bar{\bar{v}}-\frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1}}\right)$. From above, we require that $\star \in (0,1)$. Then, we have:

$$\frac{d\gamma_i}{dr_i^{\mathsf{tr}}} = \underbrace{\frac{\exp(\beta_0+\beta_r r_i^{\mathsf{tr}}+\beta_p c_i)}{f(\star)}}_{\text{positive}}\left(\underbrace{\beta_r - \frac{f'(\star)}{f(\star)}(\star)'}_{\spadesuit}\right)$$

The first term (outside the brackets) in the RHS is positive. Recall the definition of $f(\cdot)$ from (7). Using the identity that for any $z \in (0,1)$, $\frac{f'(z)}{f(z)}$ simplifies to $\frac{1}{z}+\frac{1}{(1-z)^2}$ and straightforward algebraic simplification, the term in the bracket denoted by $\spadesuit$ simplifies to:

$$\begin{aligned}
\spadesuit &= \beta_r - \underbrace{\left(\frac{1}{\star}+\frac{1}{(1-\star)^2}\right)}_{\frac{f'(\star)}{f(\star)}}\underbrace{\left(\frac{\star^2\beta_r}{(-\beta_p)k_1 v_i^{\mathsf{tr}}}\right)}_{\star'} \\
&= \underbrace{\frac{\beta_r}{(-\beta_p)k_1 v_i^{\mathsf{tr}}}}_{\text{positive}}\underbrace{\left((-\beta_p)k_1 v_i^{\mathsf{tr}} - \underbrace{\left(\star+\left(\frac{\star}{1-\star}\right)^2\right)}_{l(\star)}\right)}_{\clubsuit}.
\end{aligned}$$

The first term in the RHS is positive. Let $l(\cdot)$ denote the following:

$$l(z) = z + \left(\frac{z}{1-z}\right)^2$$

We have $l(0)=0$, $\lim_{z\uparrow 1} l(z) = \infty$, and $l(\cdot)$ is strictly increasing. Hence, for any $y>0$, $l^{-1}(y) \in (0,1)$. Consequently, we have

$$\begin{aligned}
\frac{d\gamma_i}{dr_i^{\mathsf{tr}}} > 0 \quad &\Leftrightarrow \quad \clubsuit > 0 \\
&\Leftrightarrow \quad \star < l^{-1}\left((-\beta_p)k_1 v_i^{\mathsf{tr}}\right) \\
&\Leftrightarrow \quad \frac{v_i^{\mathsf{tr}}}{l^{-1}\left((-\beta_p)k_1 v_i^{\mathsf{tr}}\right)} + \frac{\beta_r r_i^{\mathsf{tr}}}{(-\beta_p)k_1} < \bar{\bar{v}}.
\end{aligned}$$

□

A.4 Proofs of Results in Section 7

Proof of Lemma 7. Recall the solutions to q_0 in (8) and (16) under AM and PM, respectively. We will show that the RHS in (8) is larger than the RHS in (16) pointwise. Hence the fixed points, i.e., solution to q_0 satisfies $q_0^{\mathsf{AM}} > q_0^{\mathsf{PM}}$.

Consider any seller $i \in \mathcal{X}^{\mathsf{C}}$. Since $g_i(x) = f(x)e^{-h_i'^{-1}(\frac{x}{b})}$, it follows that $g_i(x) \le f(x)$ for any $x \in [bh_i'(0), 1)$, where the equality holds iff $x = bh_i'(0)$. Since $f(\cdot)$ and $g_i(\cdot)$ are monotone increasing functions (see Lemma B1), we have:

$$g_i^{-1}(y) > f^{-1}(y) \text{ for } y > bh_i'(0).$$

Using the inequality above, it follows that at any q_0, the RHS in (16) is smaller than the RHS in (8). Therefore, $q_0^{\mathsf{PM}} < q_0^{\mathsf{AM}}$.

A consequence of this result is that $\sum_{i\in[n]} q_i^{\mathsf{PM}} > \sum_{i\in[n]} q_i^{\mathsf{AM}}$, and hence $\mathcal{Q}$ is non-empty. □

Proof of Theorem 3. Since the cost of manipulation is identical across all sellers, we drop the subscript, and use $h(\cdot)$ and $g(\cdot)$.

(a) Consider $i \in \mathcal{X}^{\mathsf{C}}$. Since $x_i^{\mathsf{PM}} = 0$, from Theorem 2 we have

$$f(q_i^{\mathsf{PM}}) = A_i q_0^{\mathsf{PM}} < A_i q_0^{\mathsf{AM}} = f(q_i^{\mathsf{AM}})$$

The inequality above follows from Lemma 7 ($q_0^{\mathsf{PM}} < q_0^{\mathsf{AM}}$). Since $f(\cdot)$ is monotone, $q_i^{\mathsf{PM}} < q_i^{\mathsf{AM}}$. Since $m_i^{\mathsf{PM}} = \frac{1}{b(1-q_i^{\mathsf{PM}})}$ and $m_i^{\mathsf{AM}} = \frac{1}{b(1-q_i^{\mathsf{AM}})}$, we have $m_i^{\mathsf{PM}} < m_i^{\mathsf{AM}}$. Since $\pi_i = m_i q_i$ for $i \in \mathcal{X}^{\mathsf{C}}$, we have that $\pi_i^{\mathsf{PM}} < \pi_i^{\mathsf{AM}}$. Combining these, we have,

$$\begin{aligned} i \in \mathcal{X}^{\mathsf{C}} &\implies i \in \mathcal{Q}^{\mathsf{C}}, \quad \text{i.e.,} \quad \mathcal{X}^{\mathsf{C}} \subseteq \mathcal{Q}^{\mathsf{C}}, \text{ and} \\ i \in \mathcal{X}^{\mathsf{C}} &\implies i \in \Pi^{\mathsf{C}}, \quad \text{i.e.,} \quad \mathcal{X}^{\mathsf{C}} \subseteq \Pi^{\mathsf{C}} \end{aligned}$$

Next, consider firm $i \in \mathcal{Q}^{\mathsf{C}}$. Since $q_i^{\mathsf{AM}} > q_i^{\mathsf{PM}}$, and $m_i = \frac{1}{b(1-q_i)}$ under both AM and PM, it follows that the profit from sales, $\pi_i^{\mathsf{AM}} = m_i^{\mathsf{AM}} q_i^{\mathsf{AM}} > m_i^{\mathsf{PM}} q_i^{\mathsf{PM}} \geq m_i^{\mathsf{PM}} q_i^{\mathsf{PM}} - h(x_i^{\mathsf{PM}}) = \pi_i^{\mathsf{PM}}$. Hence, $i \in \Pi^{\mathsf{C}}$. That is, $\mathcal{Q}^{\mathsf{C}} \subseteq \Pi^{\mathsf{C}}$. Taken together,

$$\mathcal{X}^{\mathsf{C}} \subseteq \mathcal{Q}^{\mathsf{C}} \subseteq \Pi^{\mathsf{C}}.$$

(b) First, we establish that $n \in \mathcal{Q}$. Consider seller n. Recall, from Lemma 3 that $n \in \mathcal{X}$ and $\mathcal{X}$ is contiguous. Under the two settings, AM and PM, we can write q_j and q_0 in terms of q_n as follows:

$$\begin{aligned} \text{Under } \mathsf{AM}\text{: } q_j &= f^{-1}\left(\frac{A_j}{A_n} f(q_n)\right) \text{ and } q_0 = \frac{f(q_n)}{A_n} \\ \text{Under } \mathsf{PM}\text{: } q_j &= \begin{cases} g^{-1}\left(\frac{A_j}{A_n} g(q_n)\right), & \text{if } j \in \mathcal{X}; \\ f^{-1}\left(\frac{A_j}{A_n} g(q_n)\right), & \text{if } j \in \mathcal{X}^{\mathsf{C}}. \end{cases}, \text{ and } q_0 = \frac{g(q_n)}{A_n} \end{aligned}$$

The solution to q_n is obtained by solving the following:

$$\begin{aligned} q_n^{\mathsf{AM}} \text{ solves } q_n &= 1 - \left[\sum_{j \neq n} \left(\underbrace{f^{-1}\left(\frac{A_j}{A_n} f(q_n)\right)}_{q_j}\right) + \underbrace{\frac{f(q_n)}{A_n}}_{q_0}\right], \\ q_n^{\mathsf{PM}} \text{ solves } q_n &= 1 - \left[\sum_{j \neq n, j \in \mathcal{X}} \left(\underbrace{g^{-1}\left(\frac{A_j}{A_n} g(q_n)\right)}_{q_j, j \in \mathcal{X}, j \neq n}\right) + \sum_{j \in \mathcal{X}^{\mathsf{C}}} \left(\underbrace{f^{-1}\left(\frac{A_j}{A_n} g(q_n)\right)}_{q_j, j \in \mathcal{X}^{\mathsf{C}}}\right) + \underbrace{\frac{g(q_n)}{A_n}}_{q_0}\right]. \end{aligned}$$

The RHS in the equations above are $q_j, j \neq n$ and q_0 in terms of q_n. To show that $q_n^{\mathsf{PM}} > q_n^{\mathsf{AM}}$, we show that the RHS in the second equation (PM) is smaller than the RHS in the first equation (AM). It suffices to compare the terms inside the brackets.

- For $j \in \mathcal{X}, j \neq n$, the following holds:

$$f^{-1}\left(\frac{A_j}{A_n} f(q_n)\right) \geq g^{-1}\left(\frac{A_j}{A_n} g(q_n)\right).$$

  From Lemma B2 in the appendix, the above inequality follows. In particular, the inequality is strict if $A_j < A_n$. Therefore, we have the comparison for q_j, $j \in \mathcal{X}, j \neq n$.

- For $j \in \mathcal{X}^{\mathsf{C}}$, since $g(z) \leq f(z)$, we have the following:

$$f^{-1}\left(\frac{A_j}{A_n} f(q_n)\right) \geq f^{-1}\left(\frac{A_j}{A_n} g(q_n)\right)$$

- For q_0, we have $f(q_n) < g(q_n)$.

Therefore, the terms inside the bracket are higher under AM than under PM. Thus, the RHS is smaller under AM than under PM. Therefore, the solution (i.e., the fixed point) is also smaller under AM, i.e., $q_n^{\mathsf{AM}} < q_n^{\mathsf{PM}}$.

Next, we establish that $\mathcal{Q}$ is upward-closed in $[n]$. Consider seller $i \in \mathcal{X}$. Using Theorem 2, we have:

$$g(q_i^{\mathsf{PM}}) = A_i q_0^{\mathsf{PM}}.$$

Since $g(z) = f(z)e^{-h'^{-1}\left(\frac{z}{b}\right)}$, we can write the above equation as:

$$f(q_i^{\mathsf{PM}}) = A_i q_0^{\mathsf{PM}} e^{h'^{-1}\left(\frac{1}{b} g^{-1}(A_i q_0^{\mathsf{PM}})\right)}$$

Suppose $q_i^{\mathsf{PM}} > q_i^{\mathsf{AM}}$ for some $i \in [n]$. We are to show that $q_j^{\mathsf{PM}} > q_j^{\mathsf{AM}}$ for all $j > i$. Since $f(\cdot)$ is monotone, $f(q_i^{\mathsf{PM}}) > f(q_i^{\mathsf{AM}})$. Therefore,

$$\begin{aligned}
f(q_i^{\mathsf{PM}}) > f(q_i^{\mathsf{AM}}) \quad & \Leftrightarrow \quad A_i q_0^{\mathsf{PM}} e^{h'^{-1}\left(\frac{g^{-1}(A_i q_0^{\mathsf{PM}})}{b}\right)} > A_i q_0^{\mathsf{AM}} \\
& \Leftrightarrow \quad e^{h'^{-1}\left(\frac{g^{-1}(A_i q_0^{\mathsf{PM}})}{b}\right)} > \frac{q_0^{\mathsf{AM}}}{q_0^{\mathsf{PM}}} \\
& \Leftrightarrow \quad g^{-1}(A_i q_0^{\mathsf{PM}}) > b\, \underbrace{h'\left(\log \frac{q_0^{\mathsf{AM}}}{q_0^{\mathsf{PM}}}\right)}_{\triangleq \phi}. \\
& \Leftrightarrow \quad g^{-1}(A_i q_0^{\mathsf{PM}}) > b\phi \\
& \Leftrightarrow \quad A_i q_0^{\mathsf{PM}} > g(b\phi). \\
& \Leftrightarrow \quad A_i > \underbrace{\frac{1}{q_0^{\mathsf{PM}}}\left(b\phi e^{\frac{1}{1-b\phi}} e^{-h'^{-1}(\phi)}\right)}_{\text{a constant}}.
\end{aligned}$$

The RHS in the last inequality above is a constant. Thus, for any $j > i$, we have that $A_j \geq A_i \implies q_j^{\mathsf{PM}} > q_j^{\mathsf{AM}}$. Together, we have that $\mathcal{Q}$ is non-empty and upward-closed in $[n]$.

$\square$

Proof of Theorem 4. We first show the result on the seller's market share (part (a)). We then show the result on seller's profit (part (b)).

(a) From Theorem 2, recall that the equilibrium market share of seller i is as follows:

$$q_i^{\mathsf{PM}} = \begin{cases} f^{-1}(A_i q_0^{\mathsf{PM}}), & \text{if } i \in \mathcal{X}^{\mathsf{C}}; \\ g^{-1}(A_i q_0^{\mathsf{PM}}), & \text{if } i \in \mathcal{X}. \end{cases} \tag{39}$$

The following algebraic expressions are useful: Since $h(x) = \lambda(e^x - 1)$, we have:

$$\begin{aligned}
h'(x) &= h''(x) = \lambda e^x, \\
h'^{-1}(z) &= \log\left(\frac{z}{\lambda}\right).
\end{aligned}$$

Using these expressions, $g(x)$ can be written as follows:

$$g(x) = f(x)e^{-h'^{-1}\left(\frac{x}{b}\right)} = f(x)\frac{b\lambda}{x}.$$

First, we show that q_0^{PM} is increasing in λ. Since $f(\cdot)$ is increasing, it follows that q_i^{PM}, $i \in \mathcal{X}^{\mathsf{C}}$ is increasing in λ. Next, we show that q_i^{PM}, $i \in \mathcal{X}$ is decreasing in λ.
Below, we show that q_0^{PM} is decreasing in λ. Since $q_0^{\mathsf{PM}} = 1 - \sum_{i\in[n]} q_i^{\mathsf{PM}}$, we have:

$$\frac{dq_0^{\mathsf{PM}}}{d\lambda} = -\sum_{i\in[n]} \frac{dq_i^{\mathsf{PM}}}{d\lambda} \tag{40}$$

- Consider $i \in \mathcal{X}^{\mathsf{C}}$: Since $f(q_i^{\mathsf{PM}}) = A_i q_0^{\mathsf{PM}}$, differentiating both sides w.r.t. λ, we have:

$$\begin{aligned}
\frac{d}{d\lambda}\left(f(q_i)\right) &= \frac{d}{d\lambda}\left(A_i q_0^{\mathsf{PM}}\right) \\
\implies f'(q_i^{\mathsf{PM}})\frac{dq_i^{\mathsf{PM}}}{d\lambda} &= A_i \frac{dq_0^{\mathsf{PM}}}{d\lambda} \implies \frac{dq_i^{\mathsf{PM}}}{d\lambda} = \frac{1}{f'(q_i^{\mathsf{PM}})}\left(A_i \frac{dq_0^{\mathsf{PM}}}{d\lambda}\right).
\end{aligned}$$

- Consider $i \in \mathcal{X}$: Since $g(q_i^{\mathsf{PM}}) = A_i q_0^{\mathsf{PM}}$,

$$\begin{aligned}
\frac{d}{d\lambda}\left(g(q_i^{\mathsf{PM}})\right) &= \frac{d}{d\lambda}\left(A_i q_0^{\mathsf{PM}}\right) \\
\implies g'(q_i^{\mathsf{PM}})\frac{dq_i^{\mathsf{PM}}}{d\lambda} + \frac{\partial g(q_i^{\mathsf{PM}})}{\partial \lambda} &= A_i \frac{dq_0^{\mathsf{PM}}}{d\lambda} \\
\implies \frac{dq_i^{\mathsf{PM}}}{d\lambda} &= \frac{1}{g'(q_i^{\mathsf{PM}})}\left(A_i \frac{dq_0^{\mathsf{PM}}}{d\lambda} - \frac{\partial g(q_i^{\mathsf{PM}})}{\partial \lambda}\right).
\end{aligned}$$

Substituting the above in the RHS of (40) and using the identities, we have:

$$\frac{dq_i^{\mathsf{PM}}}{d\lambda} = \begin{cases} \frac{A_i}{f'(q_i^{\mathsf{PM}})}\frac{dq_0^{\mathsf{PM}}}{d\lambda}, & \text{if } i \in \mathcal{X}^{\mathsf{C}}; \\ \frac{1}{g'(q_i^{\mathsf{PM}})}\left(A_i \frac{dq_0^{\mathsf{PM}}}{d\lambda} - \frac{\partial g(q_i^{\mathsf{PM}})}{\partial \lambda}\right), & \text{if } i \in \mathcal{X}. \end{cases}$$

In the RHS above, using algebraic manipulation, we have $\frac{\partial g(q_i^{\mathsf{PM}})}{\partial \lambda} = \frac{g(q_i^{\mathsf{PM}})}{\lambda} = \frac{A_i q_0^{\mathsf{PM}}}{\lambda}$. Substituting the above in (40), we have:

$$\begin{aligned}
\frac{dq_0^{\mathsf{PM}}}{d\lambda} &= \frac{\sum_{i\in\mathcal{X}} \frac{\frac{\partial g(q_i^{\mathsf{PM}})}{\partial \lambda}}{g'(q_i^{\mathsf{PM}})}}{1 + \sum_{i\in\mathcal{X}^{\mathsf{C}}} \frac{A_i}{f'(q_i^{\mathsf{PM}})} + \sum_{i\in\mathcal{X}} \frac{A_i}{g'(q_i^{\mathsf{PM}})}} \\
&= \frac{q_0^{\mathsf{PM}}\tau}{\lambda}.
\end{aligned} \tag{41}$$

Recall the definition of τ in (**??**). Indeed, $\tau \in (0, 1)$. The RHS of (41) is positive. Therefore, q_0^{PM} is increasing in λ. Consequently, q_i^{PM}, $i \in \mathcal{X}^{\mathsf{C}}$ is also increasing in λ.
Now, consider q_i^{PM}, $i \in \mathcal{X}$. Using (39), we have:

$$\begin{aligned}
\frac{dq_i^{\mathsf{PM}}}{d\lambda} &= \frac{1}{g'(q_i^{\mathsf{PM}})}\left(\underbrace{A_i \frac{dq_0^{\mathsf{PM}}}{d\lambda} - \frac{\partial g(q_i^{\mathsf{PM}})}{\partial \lambda}}_{<0}\right) \\
&= -\frac{1}{g'(q_i^{\mathsf{PM}})}\frac{q_0^{\mathsf{PM}}}{\lambda}(1-\tau).
\end{aligned} \tag{42}$$

Since $\tau \in (0, 1)$, the RHS above is negative. Therefore, q_i^{PM}, $i \in \mathcal{X}$, is decreasing in λ.

(b) To show that π_i^{PM}, $i \in [n]$, is increasing in λ, we first show the result for $i \in \mathcal{X}^{\mathsf{C}}$. We then show the result for $i \in \mathcal{X}$. Consider seller $i \in \mathcal{X}^{\mathsf{C}}$. Since $x_i = 0$, seller i's profit is

$$\pi_i = m_i q_i = \frac{q_i^{\mathsf{PM}}}{b(1-q_i^{\mathsf{PM}})},$$

which is monotone in q_i^{PM}. From part (a) of this result, since q_i^{PM} is increasing in λ, it follows that π_i^{PM} is increasing in λ. Now, consider seller $i \in \mathcal{X}$. Recall that seller i's equilibrium profit is

$$\pi_i = \underbrace{m_i q_i}_{\text{direct profit from sales}} - \underbrace{h\left(h'^{-1}\left(\frac{q_i}{b}\right)\right)}_{\text{cost of manipulation}}.$$

Substituting for $h(\cdot)$ and $h'^{-1}(\cdot)$, we have:

$$\begin{aligned}
\pi_i^{\mathsf{PM}} &= \frac{q_i^{\mathsf{PM}}}{b(1-q_i^{\mathsf{PM}})} - \lambda\left(\frac{q_i^{\mathsf{PM}}}{b\lambda} - 1\right) \\
\implies \frac{d\pi_i^{\mathsf{PM}}}{d\lambda} &= \frac{(2-q_i)q_i}{b(1-q_i)^2}\frac{dq_i}{d\lambda} + 1.
\end{aligned}$$

For the RHS to be positive, we require the following condition to hold:

$$1 > \left(-\frac{dq_i^{\mathsf{PM}}}{d\lambda}\right)\left(\frac{(2-q_i^{\mathsf{PM}})q_i^{\mathsf{PM}}}{b(1-q_i^{\mathsf{PM}})^2}\right). \tag{43}$$

Substituting for $\frac{dq_i^{\mathsf{PM}}}{d\lambda}$ from (42), the condition in (43) simplifies to:

$$\lambda > \frac{q_i^{\mathsf{PM}}(2-q_i^{\mathsf{PM}})}{b}\,(1-\tau)\,.$$

In the RHS, $q_i(2-q_i)$ is increasing in $q_i \in [0,1]$. If $A_1 \le A_2 \le \ldots A_n$, and hence, q_i is an increasing sequence, an upper bound on the RHS is $\frac{q_n^{\mathsf{PM}}(2-q_n^{\mathsf{PM}})}{b}\,(1-\tau)$

□

Proof of Theorem 5. Consider the consumer utility model below

$$u_i = a_i + \delta x_i - bp_i + \epsilon_i$$

Recall the definition of γ_i in (15):

$$\gamma_i = \begin{cases} \frac{A_i}{f\left(bh_i'(0)\right)}, & \text{if } h_i'(0) < \frac{1}{b}; \\ 0, & \text{otherwise.} \end{cases}$$

Let $\hat{x}_i = \delta x_i$. We can rewrite the game in this setting as a game where firms choose $(\hat{x}_i, p_i)$, and face the following cost of manipulation $\hat{h}_i^{(\delta)}(\cdot)$:

$$\hat{h}_i^{(\delta)}(x) = h_i\left(\frac{x}{\delta}\right)$$

Fix δ. In this game, we have:

$$\gamma_i^{(\delta)} = \begin{cases} \frac{A_i}{f\left(\frac{b}{\delta}h_i'(0)\right)}, & \text{if } h_i'(0) < \frac{\delta}{b}; \\ 0, & \text{otherwise.} \end{cases}$$

Observe that $\gamma_i(\delta)$ is (weakly) increasing in δ. Next, we will show that $q_0^{\mathsf{PM}(\delta)}$ is decreasing in δ. Let $g_i^{(\delta)}(z)$ denote the following:

$$g_i^{(\delta)}(z) = ze^{\frac{1}{1-z}}e^{-\delta h_i'^{-1}\left(\frac{\delta}{b}z\right)} = \frac{f(z)}{e^{\delta h_i'^{-1}\left(\frac{\delta}{b}z\right)}}.$$

Observe that $g_i^{(\delta)}(z)$ is decreasing in δ. For fixed δ, $g_i^{(\delta)}(\cdot)$ is a monotone increasing function. We can show that q_0^{PM} is the solution to the following fixed point equation:

$$q_0 = 1 - \sum_{i\in\mathcal{X}} g_i^{(\delta)^{-1}}(A_i q_0) - \sum_{i\in\mathcal{X}^{\mathsf{C}}} f^{-1}(A_i q_0).$$

For any δ, following the same approach as in Theorem 2, a solution to q_0 exists and is unique. Let $q_0^{\mathsf{PM}(\delta)}$ denote this solution. Since $g_i^{(\delta)}(z)$ is decreasing in δ and $g_i^{(\delta)}(\cdot)$ is a monotone increasing function, it follows that $q_0^{\mathsf{PM}(\delta)}$ is decreasing in δ. Combining the observations that $\gamma_i{}^{(\delta)}$ is increasing in δ and $q_0^{\mathsf{PM}(\delta)}$ is decreasing in δ, we have the result that for any $\delta \le \delta'$, $\mathcal{X}^{(\delta)} \subseteq \mathcal{X}^{(\delta')}$. □

Proof of Theorem 6. We simplify the industry revenues using seller i's best response in (10):

$$\begin{aligned}
\sum_i p_i q_i &= \sum_i \left(c_i + \underbrace{m_i}_{=\frac{1}{b(1-q_i)}}\right) q_i \\
&= \sum_i c_i q_i + \frac{1}{b}\sum_i \underbrace{\frac{q_i}{1-q_i}}_{\frac{1}{1-q_i}-1}. \\
&= \sum_i c_i q_i - \frac{n}{b} + \frac{1}{b}\sum_i \frac{1}{1-q_i}.
\end{aligned}$$

The RHS comprises of three terms. The second term is a constant. Consider the first term. Recall from Lemma 7 that $\sum_{i\in[n]} q_i^{\mathsf{PM}} > \sum_{i\in[n]} q_i^{\mathsf{AM}}$. Denote the following:

$$\begin{aligned}
\Delta_1 &= \sum_{i\in\mathcal{Q}}(q_i^{\mathsf{PM}} - q_i^{\mathsf{AM}}),\\
\Delta_2 &= \sum_{i\in\mathcal{Q}^{\mathsf{C}}}(q_i^{\mathsf{AM}} - q_i^{\mathsf{PM}}),\\
\bar{i}_{\mathcal{Q}^{\mathsf{C}}} &= \max_{i\in\mathcal{Q}^{\mathsf{C}}} i, \underline{i}_{\mathcal{Q}} = \min_{i\in\mathcal{Q}} i.
\end{aligned}$$

From Lemma 7, it holds that $\Delta_1 > \Delta_2$. Next,

$$\begin{aligned}
\sum_{i\in\mathcal{Q}} c_i(q_i^{\mathsf{PM}} - q_i^{\mathsf{AM}}) &\geq c_{\underline{i}_{\mathcal{Q}}}\Delta_1\\
&> c_{\bar{i}_{\mathcal{Q}^{\mathsf{C}}}}\Delta_2\\
&\geq \sum_{i\in\mathcal{Q}^{\mathsf{C}}} c_i(q_i^{\mathsf{AM}} - q_i^{\mathsf{PM}}).
\end{aligned}$$

The first and third inequalities above follow from $c_1 \leq c_2 \leq \ldots c_n$, and the second inequality follows from the observation that $\bar{i}_{\mathcal{Q}^{\mathsf{C}}} < \underline{i}_{\mathcal{Q}}$. Therefore, $\sum_{i\in[n]} c_i(q_i^{\mathsf{PM}} - q_i^{\mathsf{AM}}) > 0$. Consider the third term. Define the function

$$s(z) = \frac{1}{1-z}.$$

$s(z)$ is strictly convex and strictly increasing in z. For any $\mathbf{q} = (q_1, q_2, \ldots, q_n)$, where $q_1 \leq q_2 \leq \ldots q_n$, define the following function:

$$S(\mathbf{q}) = \sum_{i\in[n]} s(q_i).$$

We compare $S(\mathbf{q}^{\mathsf{AM}})$ and $S(\mathbf{q}^{\mathsf{PM}})$. Depending on whether $\mathcal{Q}^{\mathsf{C}}$ is empty or non-empty, we have the following cases.

- Suppose $\mathcal{Q}^{\mathsf{C}}$ is empty. Then, $q_i^{\mathsf{PM}} > q_i^{\mathsf{AM}}$ for all i. Since $s(z)$ is increasing in z, it follows that $s(q_i^{\mathsf{PM}}) > s(q_i^{\mathsf{AM}})$ for all i. Thus, $S(\mathbf{q}^{\mathsf{PM}}) > S(\mathbf{q}^{\mathsf{AM}})$.
- Suppose $\mathcal{Q}^{\mathsf{C}}$ is non-empty. Recall, from Theorem 3, that $n \in \mathcal{Q}$. We construct a lower bound on $S(\mathbf{q}^{\mathsf{PM}}) - S(\mathbf{q}^{\mathsf{AM}})$, and show that the lower bound is positive. Therefore,

  $$S(\mathbf{q}^{\mathsf{PM}}) - S(\mathbf{q}^{\mathsf{AM}}) = \sum_{i\in\mathcal{Q}}\left(s(q_i^{\mathsf{PM}}) - s(q_i^{\mathsf{AM}})\right) - \sum_{i\in\mathcal{Q}^{\mathsf{C}}}\left(s(q_i^{\mathsf{AM}}) - s(q_i^{\mathsf{PM}})\right)$$

  Due to the convexity of $s(z)$, we have:

  $$\begin{aligned}
  \sum_{i\in\mathcal{Q}}\left(s(q_i^{\mathsf{PM}}) - s(q_i^{\mathsf{AM}})\right) &\geq s'(q_{\underline{i}_{\mathcal{Q}}}^{\mathsf{AM}})\Delta_1\\
  \sum_{i\in\mathcal{Q}^{\mathsf{C}}}\left(s(q_i^{\mathsf{AM}}) - s(q_i^{\mathsf{PM}})\right) &\leq s'(q_{\bar{i}_{\mathcal{Q}^{\mathsf{C}}}}^{\mathsf{AM}})\Delta_2.
  \end{aligned}$$

  Due to convexity of $s(z)$ (i.e., $s'(z)$ is increasing in z) and the observation that q_i^{AM} is increasing in A_i (in Theorem 1), we have:

  $$s'(q_{\underline{i}_{\mathcal{Q}}}^{\mathsf{AM}}) \geq s'(q_{\bar{i}_{\mathcal{Q}^{\mathsf{C}}}}^{\mathsf{AM}}).$$

  Since $\Delta_1 > \Delta_2$, we have:

  $$\begin{aligned}
  s'(q_{\underline{i}_{\mathcal{Q}}}^{\mathsf{AM}})\Delta_1 > s'(q_{\bar{i}_{\mathcal{Q}^{\mathsf{C}}}}^{\mathsf{AM}})\Delta_2 &\implies \sum_{i\in\mathcal{Q}}\left(s(q_i^{\mathsf{PM}}) - s(q_i^{\mathsf{AM}})\right) > \sum_{i\in\mathcal{Q}^{\mathsf{C}}}\left(s(q_i^{\mathsf{AM}}) - s(q_i^{\mathsf{PM}})\right)\\
  &\implies S(\mathbf{q}^{\mathsf{PM}}) > S(\mathbf{q}^{\mathsf{AM}}).
  \end{aligned}$$

$\square$

B Helpful Results

Some helpful results and their proofs are listed below.

Lemma B1. *Suppose $h_i'(0) < \frac{1}{b}$. Under Assumption 1, $g_i(z)$ is increasing in $z \in [bh_i'(0), 1)$, $g_i(bh_i'(0)) = f(bh'(0)) = bh'(0)e^{\frac{1}{1-bh_i'(0)}}$, $g_i(1) = \infty$.*

Proof of Lemma B1. For any $z \in [bh_i'(0), 1)$,

$$g'(z) \quad = \quad e^{\frac{1}{1-z} - h_i'^{-1}(\frac{1}{b}z)} \left(1 - \frac{\frac{z}{b}}{h_i''(h_i'^{-1}(\frac{z}{b}))} + \frac{z}{(1-z)^2}\right)$$

The term outside the brackets in the RHS is positive. We show that the sum of the first and second term inside the bracket is positive. Substituting $z \to yb$, it suffices to show that $1 - \frac{y}{h_i''(h_i'^{-1}(y))} > 0$, where $h_i'(0) \leq y < \frac{1}{b}$. Since $h_i(\cdot)$ is convex, it suffices to show $y < h_i''(h_i'^{-1}(y))$. Substitute $\hat{y} = h_i'^{-1}(y)$. Then, we are required to show $h_i'(\hat{y}) < h_i''(\hat{y})$, which holds from Assumption 1(b). ☐

Lemma B2. *Fix $k \in (0, 1]$. For any $z \in (0, 1)$, $g_i^{-1}(kg_i(z)) \leq f^{-1}(kf(z))$, where equality holds iff $k = 1$.*

Proof of Lemma B2. For convenience, we drop the subscript i. Let

$$\begin{aligned} z_1 &= g^{-1}(kg(z)) \implies g(z_1) = kg(z) \\ z_2 &= f^{-1}(kf(z)) \implies f(z_2) = kf(z). \end{aligned}$$

We are required to show that $z_1 < z_2$. Combining these two equations and using (14), we have:

$$\frac{f(z_1)}{f(z_2)} = e^{\left(h'^{-1}(\frac{z_1}{b}) - h'^{-1}(\frac{z}{b})\right)}.$$

Since $k \in (0, 1]$, we have that $z_1 \leq z$ with a strict inequality if $k < 1$. Therefore, the RHS is less than or equal to 1. Since $f(\cdot)$ is monotone and positive, $z_1 \leq z_2$. ☐

Recall the definition of $r(q)$ from (**??**) in Section 7:

$$r(q) = q(2-q)\left(1 - \frac{2(1-q)^2}{2q^2 - 6q + 3}\right).$$

Lemma B3. *For $0 \leq q < \frac{1}{2}$, we have :*

(a) $r(q)$ is concave in q.

(b) $r(q) < \frac{1}{2}$.

Proof of Lemma B3. (a) We express $r(q)$ as follows:

$$r(q) = \frac{q(2-q)(1-2q)}{2q^2 - 6q + 3}$$

To show that $r(q)$ is concave in $q \in \left[0, \frac{1}{2}\right]$, we will show that $r''(q) < 0$ in $q \in \left[0, \frac{1}{2}\right]$. $r''(q)$ is as follows:

$$r''(q) = \frac{\overbrace{2(8q^3 - 18q^2 + 18q - 9)}^{P(q),\ \text{negative in } [0,\, \frac{1}{2}]}}{\underbrace{\left(2\left(\frac{3}{2} - q\right)^2 - \frac{3}{2}\right)^3}_{\text{positive in } [0,\, \frac{1}{2}]}}$$

The denominator in the RHS above is decreasing in q for $q \in \left[0, \frac{1}{2}\right]$ and is positive at $q = \frac{1}{2}$. Therefore, the denominator is positive for $q \in \left[0, \frac{1}{2}\right]$. For convenience, denote the numerator in the RHS by $P(q)$. We show that $P(q) < 0$ for $q \in \left[0, \frac{1}{2}\right]$. We have:

$$P'(q) = 3\left(16\left(q - \frac{3}{4}\right)^2 + \frac{3}{4}\right)$$

Observe that $P'(q) > 0$. Therefore, $P(q)$ is increasing in q for $q \in \left[0, \frac{1}{2}\right]$. Through straightforward algebra, we can verify that $P\left(\frac{1}{2}\right) = -7$. Therefore, $P(q) < 0$ for $q \in \left[0, \frac{1}{2}\right]$. So, we can conclude that $r''(q) < 0$ for $q \in \left[0, \frac{1}{2}\right]$.

(b) The first derivative of $r(q)$ is as follows:

$$r'(q) = \frac{2(1-q)}{(2q^2 - 6q + 3)^2} \underbrace{(-2q^3 + 10q^2 - 12q + 3)}_{\triangleq M(q)} \tag{44}$$

From part (a) of this result, FOC's are necessary and sufficient to find a global maximizer of $r(q)$ for $q \in [0, 0.5]$. Setting the RHS of (44) to 0 is equivalent to setting the last term in the RHS above, denoted by $M(q)$, to 0. Let Δ as the discriminant of $M(q)$; $\Delta = 564$. This implies that we have three distinct real roots for $M(q) = 0$. Let $q_{(i)}$ denote the i^{th} root for $M(q) = 0$. We have:

$$\begin{aligned}
q_{(1)} &= \frac{5}{3} + \frac{7}{9\left(-\frac{1}{2} + \frac{\sqrt{3}i}{2}\right)\sqrt[3]{\frac{41}{108} + \frac{\sqrt{47}i}{12}}} + \left(-\frac{1}{2} + \frac{\sqrt{3}i}{2}\right)\sqrt[3]{\frac{41}{108} + \frac{\sqrt{47}i}{12}} \\
&\approx 0.33956 \\
q_{(2)} &= \frac{5}{3} + \left(-\frac{1}{2} - \frac{\sqrt{3}i}{2}\right)\sqrt[3]{\frac{41}{108} + \frac{\sqrt{47}i}{12}} + \frac{7}{9\left(-\frac{1}{2} - \frac{\sqrt{3}i}{2}\right)\sqrt[3]{\frac{41}{108} + \frac{\sqrt{47}i}{12}}} \\
&\approx 1.3240 \\
q_{(3)} &= \frac{5}{3} + \frac{7}{9\sqrt[3]{\frac{41}{108} + \frac{\sqrt{47}i}{12}}} + \sqrt[3]{\frac{41}{108} + \frac{\sqrt{47}i}{12}} \\
&\approx 3.3364
\end{aligned}$$

Correspondingly, $r(q_{(1)}) \approx 0.152$.

$\square$

C Maximum Likelihood Estimation of the MNL Choice Model with Unobserved No Purchase Data

Consider the maximum likelihood estimation procedure where the consumer utility model is as follows:

$$u_i = \beta_0 + \beta_p p_i + \beta_r r_i + \epsilon_i,\ i \in [n],\ u_0 = 0,$$

ϵ_i are i.i.d. Gumbel random variables, and consumer purchase model follows

$$q_i = \frac{e^{u_i}}{1 + \sum_{j \in [K]} e^{u_j}},\ i \in [n] \cup \{0\}.$$

The likelihood function, given purchase and no-purchase data ($\mathbf{s} = (s_i)_{i \in [n]}, s_0$), is as follows:

$$\mathcal{L}(\boldsymbol{\beta}|\mathbf{s}, s_0) = \frac{\overline{s}!}{\prod_{i \in [n] \cup \{0\}} s_i!} \prod_{i \in [n] \cup \{0\}} q_i^{s_i}.$$

where $\overline{s} = \sum_{i \in [n] \cup \{0\}} s_i$. The log-likelihood is as follows:

$$\log \mathcal{L}(\boldsymbol{\beta}|\mathbf{n}, n_0) = \sum_{i \in [n]} s_i \log q_i + s_0 \log q_0 + \mathsf{SL}(\overline{s}) - \sum_{i \in [n] \cup \{0\}} \mathsf{SL}(s_i),$$

where, for any $j \in \mathbb{I}^+$, $\mathsf{SL}(j) = \sum_{j'=1}^{j} \log(j')$.

Let $\hat{q}_i$ denote the empirical market share, i.e., $\hat{q}_i = \frac{s_i}{\overline{s}}$. Corresponding to any vector $\mathbf{y} = \{y_i\}_{i \in [n]}$, consider the random variable:

$$\tilde{y} = q_0 \circ 0 + \sum_{i \in [n]} q_i \circ y_i.$$

Define the following operators:

$$\begin{aligned}
\hat{\mathbb{E}}[\mathbf{y}] &= \mathbb{E}[\tilde{y}] = \sum_{i \in [n]} q_i y_i, \\
\hat{\mathrm{Var}}[\mathbf{y}] &= \mathrm{Var}[\tilde{y}] = \sum_{i \in [n]} q_i y_i^2 - \left(\sum_{i \in [K]} q_i y_i \right)^2, \text{ and} \\
\hat{\mathrm{Cov}}[\mathbf{y}_1, \mathbf{y}_2] &= \mathbb{E}[\tilde{y}_A \tilde{y}_B] - \mathbb{E}[\tilde{y}_A]\mathbb{E}[\tilde{y}_B].
\end{aligned}$$

We have the following:

$$\begin{aligned}
\nabla \log \mathcal{L}(\boldsymbol{\beta}|\mathbf{s}, s_0) &= \overline{n} \left[\sum_{i \in [n]} (\hat{q}_i - q_i) \quad \sum_{i \in [n]} p_i (\hat{q}_i - q_i) \quad \sum_{i \in [n]} r_i (\hat{q}_i - q_i) \right] \\
\text{and } H \triangleq \nabla^2 \log \mathcal{L}(\boldsymbol{\theta}|\mathbf{s}, s_0) &= -\overline{n} \begin{bmatrix} q_0 \sum_j q_j & q_0 \hat{\mathbb{E}}[\mathbf{p}] & q_0 \hat{\mathbb{E}}[\mathbf{r}] \\ q_0 \hat{\mathbb{E}}[\mathbf{p}] & \hat{\mathrm{Var}}[\mathbf{p}] & \hat{\mathrm{Cov}}[\mathbf{p}, \mathbf{r}] \\ q_0 \hat{\mathbb{E}}[\mathbf{r}] & \hat{\mathrm{Cov}}[\mathbf{p}, \mathbf{r}] & \hat{\mathrm{Var}}[\mathbf{r}] \end{bmatrix}
\end{aligned}$$

For given s_0, H is negative semi-definite.[10] Hence, FOC's identify the MLE of $\boldsymbol{\beta}$. Since s_0 is unknown, we identify the maximum likelihood estimates for $\boldsymbol{\beta}$, denoted by $\boldsymbol{\beta}^{\mathsf{MLE}}$ using the EM approach.

Input: $\mathbf{s}, \mathbf{p}, \mathbf{r}, n, \varepsilon$
$\boldsymbol{\beta}_0 \leftarrow \mathbf{0}$, $\omega \leftarrow 0$
repeat
    $\omega \leftarrow \omega + 1$ ▷ E-Step

[10] This follows from the fact that the variance-covariance matrix of a multivariate distribution is positive definite.

for $i \in [n]$ **do**
  $u_i \leftarrow \beta_0^\omega + \beta_p^\omega p_i + \beta_r^\omega r_i$
  $q_i \leftarrow \frac{e^{u_i}}{1+\sum_{j\in[n]} e^{u_j}}$
  $s_0 \leftarrow \frac{q_0}{1-q_0}\overline{s}$
end for ▷ M-Step
$\boldsymbol{\beta}_\omega \leftarrow \arg\max_{\boldsymbol{\beta}\in\Re^3} \log \mathcal{L}(\boldsymbol{\beta}|\mathbf{s}, \mathbf{s}_0)$
until $\|\boldsymbol{\beta}_\omega - \boldsymbol{\beta}_{\omega-1}\| < \epsilon$

We now substitute $\boldsymbol{\beta}^{\mathsf{MLE}}$ (obtained from the above approach) in H. In particular, $\mathbf{q} \equiv \mathbf{q}(\boldsymbol{\beta}^{\mathsf{MLE}})$. We identify the inverse of H, denoted by H^{-1}. The standard errors of $\boldsymbol{\beta}^{\mathsf{MLE}}$ corresponds to the diagonal elements of H^{-1}.

In our model, we have data across 24 weeks for 11 products. We pool the data across the 24 weeks to estimate $\boldsymbol{\beta}$. Our estimation results are shown in Table 1.

D Consumer Surplus in the MNL Model

Consider the canonical MNL choice model, where the utility from consuming product $i \in [n]$, and 0 (the no-purchase option) is:

$$\begin{aligned} u_0 &= \epsilon_0, \\ u_i &= v_i + \epsilon_i \text{ for } i \in [n] \end{aligned}$$

where ϵ_0 and ϵ_i, $i \in [n]$ are i.i.d Gumbel r.v.'s. A representative consumer chooses the no purchase option or one of the products $i \in [n]$ as follows:

$$\begin{aligned} \text{consumer chooses } i \text{ if: } v_i + \epsilon_i &> \max\{0, \max_{j \neq i}\{v_j + \epsilon_j\}\} \\ \text{consumer chooses 0 if: } \epsilon_0 &> \max_{i \in [n]}\{v_i + \epsilon_i\}. \end{aligned}$$

Recall that the consumer purchases product i w.p. q_i as follows:

$$\begin{aligned} q_0 &= \frac{1}{1 + \sum_{j \in [n]} e^{v_j}} \\ q_i &= \frac{e^{v_i}}{1 + \sum_{j \in [n]} e^{v_j}} \text{ for } i \in [n] \end{aligned}$$

That is, the consumer purchase follows the following discrete distribution over $[n] \cup \{0\}$:

$$\text{Consumer Purchase } \sim q_0 \circ 0 + \sum_{i \in [n]} q_i \circ i.$$

The expected consumer surplus (where the expectations are taken over $\boldsymbol{\epsilon}$) is as follows:

$$\begin{aligned} \text{Expected CS} &= q_0 \mathbb{E}_{\epsilon_0, \boldsymbol{\epsilon}}[\epsilon_0 | \epsilon_0 > \max_{i \in [n]}(v_i + \epsilon_i)] + \\ &\quad \sum_{i \in [n]} q_i \mathbb{E}_{\epsilon_0, \boldsymbol{\epsilon}}[(v_i + \epsilon_i) | v_i + \epsilon_i > \max\{\epsilon_0, \max_{j \in [n] \setminus \{i\}}(v_j + \epsilon_j)\}] \end{aligned}$$

Recall that

$$\begin{aligned} q_0 &= \mathbb{E}_{\epsilon_0, \boldsymbol{\epsilon}}\left[\mathbf{1}\left(\epsilon_0 | \epsilon_0 > \max_{i \in [n]} v_i + \epsilon_i\right)\right] \\ q_i &= \mathbb{E}_{\epsilon_0, \boldsymbol{\epsilon}}\left[\mathbf{1}\left(v_i + \epsilon_i > \max\{\epsilon_0, \max_{j \in [n], j \neq i} v_j + \epsilon_j\}\right)\right]. \end{aligned}$$

$$\begin{aligned} &= \mathbb{E}_{\boldsymbol{\epsilon}}\left[\int_{y_0 \in \Re} y_0 \mathbf{1}\left(y_0 > \max_{i \in [n]}(v_i + \epsilon_i)\right) f(y_0) dy_0\right] + \\ &\quad \left[\sum_{i \in [n]} \mathbb{E}_{\epsilon_0, \boldsymbol{\epsilon}_{-i}} \int_{y_i \in \Re} (v_i + y_i) \mathbf{1}\left(v_i + y_i > \max\{\epsilon_0, \max_{j \in [n] \setminus \{i\}}(v_j + \epsilon_j)\}\right) f(y_i) dy_i\right] \\ &= \int_{y_0 \in \Re} y_0 \prod_{i \in [n]} F(y_0 - v_i) f(y_0) dy_0 + \\ &\quad \sum_{i \in [n]} \int_{y_i \in \Re} (v_i + y_i) F(v_i + y_i) \prod_{j \in [n] \setminus \{i\}} F(v_i - v_j + y_i) f(y_i) dy_i. \end{aligned}$$

where $f(\cdot)$ and $F(\cdot)$ represent the p.d.f. and c.d.f. of the standard Gumbel distribution. Expanding the above,

$$\begin{aligned}\text{Expected CS} &= \int_{y_0\in\Re} y_0 e^{-(y_0+\sum_{j\in[n]} e^{-y_0}\beta_j^0)} dy_0 \\ &\quad + \sum_{i\in[n]} \left(v_i q_i + \int_{y_i\in\Re} y_i e^{-(y_i+\sum_{j\neq i} e^{-y_i}\beta_j^i + e^{-y_i}\beta_0^i)} dy_i \right)\end{aligned}$$

where $\beta_j^0 = e^{v_j - v_0} = e^{v_j}$ and $\beta_j^i = e^{v_j - v_i}$.

$$\begin{aligned}\text{Expected CS} &= q_0\left(\gamma - \log(q_0)\right) + \sum_{i\in[n]} q_i\left(\gamma - \log(q_0)\right) \\ &= \underbrace{\gamma}_{\text{Euler's constant, } \approx 0.577} \underbrace{- \log(q_0)}_{=\log\left(1+\sum_{i\in[n]} e^{v_i}\right)} .\end{aligned}$$

In the absence of manipulation, observe that

$$\begin{aligned}v_0^{\mathsf{AM}} &= 0 \\ v_i^{\mathsf{AM}} &= a_i - bp_i^{\mathsf{AM}} = \log(A_i) - bm_i^{\mathsf{AM}}.\end{aligned}$$

Therefore,

$$\text{Expected CS under } \mathsf{AM} = \gamma - \log(q_0^{\mathsf{AM}}).$$

In the presence of manipulation, prior to purchase, consumers' anticipate the following:

$$v_i^{\mathsf{PM}} = a_i + x_i^{\mathsf{PM}} - bp_i^{\mathsf{PM}} = \log(A_i) + x_i^{\mathsf{PM}} - bm_i^{\mathsf{PM}}.$$

Indeed, the market shares q_i^{PM} result from the above values of v_i. However, post-purchase, consumers realize the *true* utility as follows:

$$\tilde{v}_i^{\mathsf{PM}} = a_i - bp_i^{\mathsf{PM}} = \log(A_i) - bm_i^{\mathsf{PM}}.$$

The post-purchase consumer surplus can be written as follows:

$$\begin{aligned}\text{Consumer Surplus} &= \underbrace{\mathbb{E}_{\boldsymbol{\epsilon}}\left[\int_{y_0\in\Re} y_0 \mathbf{1}\left(y_0 > \max_{i\in[n]}(v_i^{\mathsf{PM}} + \epsilon_i)\right) f(y_0)dy_0\right]}_{CS_0} + \\ &\quad \sum_{i\in[n]} \underbrace{\mathbb{E}_{\epsilon_0,\boldsymbol{\epsilon}_{-i}}\left[\int_{y_i\in\Re} (\tilde{v}_i^{\mathsf{PM}} + y_i)\mathbf{1}\left(v_i^{\mathsf{PM}} + y_i > \max\{\epsilon_0, \max_{j\in[n], j\neq i}(v_j^{\mathsf{PM}} + \epsilon_j)\}\right) f(y_i)dy_i\right]}_{CS_i}\end{aligned}$$

In the second term, add and subtract x_i to $\tilde{v}_i^{\mathsf{PM}} + y_i$. We have:

$$\text{Consumer Surplus} = \gamma - \log\left(q_0^{\mathsf{PM}}\right) - \underbrace{\sum_{i\in[n]} x_i^{\mathsf{PM}} q_i^{\mathsf{PM}}}_{=\overline{x}^{\mathsf{PM}}, \text{ the average level of manipulation}} .$$

Therefore,

$$CS^{\mathsf{AM}} > CS^{\mathsf{PM}} \Leftrightarrow \overline{x}^{\mathsf{PM}} > \log\left(\frac{q_0^{\mathsf{AM}}}{q_0^{\mathsf{PM}}}\right).$$

E Verification: Equation (21) Satisfies Assumption 1

For notational convenience, we denote $\tilde{x_i} = \frac{x_i}{\beta_r}$, and $y_i = v_i^{\mathsf{tr}} \frac{\tilde{x_i}}{R - r_i^{\mathsf{tr}} - \tilde{x_i}}$. Then, we re-write (21) as

$$h_i\left(y(\tilde{x_i})\right) = k_1 y_i(\tilde{x_i}) + k_2 y_i^2(\tilde{x_i}). \tag{45}$$

To verify Assumption 1(a), we derive $h_i'(x_i)$ and $h_i''(x_i)$ as

$$\begin{aligned} h_i'(x_i) = &\quad k_1 \frac{dy_i}{dx_i} + 2k_2 y_i(\tilde{x_i}) \frac{dy_i}{dx_i} \\ = &\quad \frac{dy_i}{dx_i}\left(k_1 + 2k_2 y_i(\tilde{x_i})\right) \end{aligned} \tag{46}$$

$$h_i''(x_i) = \frac{d^2 y_i}{dx_i^2}\left(k_1 + k_2 y_i(\tilde{x_i})\right) + 2k_2 \left(\frac{dy_i}{dx_i}\right)^2 \tag{47}$$

By direct calculation, we have:

$$\begin{aligned} \frac{dy_i}{dx_i} &= \frac{1}{\beta_r}\frac{dy_i}{d\tilde{x_i}} \\ &= \frac{v_i^{\mathsf{tr}}(R - r_i^{\mathsf{tr}})}{\beta_r(R - r_i^{\mathsf{tr}} - \tilde{x_i})^2} \\ \frac{d^2 y_i}{dx_i^2} &= \frac{2v_i^{\mathsf{tr}}(R - r_i^{\mathsf{tr}})}{\beta_r^2(R - r_i^{\mathsf{tr}} - \tilde{x_i})^3} \end{aligned} \tag{48}$$

Since $\tilde{x_i} \in [0, R - r_i^{\mathsf{tr}})$ and $r_i^{\mathsf{tr}} \in [0, R)$, Assumption 1$(a)$ is satisfied. Next, we verify Assumption 1(b) by showing that $\frac{h_i''(x_i)}{h_i'(x_i)} \geq 1$. Recall from (46), (47) and (48), we have:

$$\begin{aligned} \frac{h_i''}{h_i'} &= \frac{y_i''(k_1 + k_2 y_i) + 2k_2(y_i')^2}{y_i'(k_1 + 2k_2 y_i)} \\ &= \frac{y_i''}{y_i'} + \frac{2k_2 y_i'}{k_1 + 2k_2 y_i} \\ &\geq 2\sqrt{\frac{2k_2 y_i''}{k_1 + 2k_2 y_i}} \\ &= 2\sqrt{\frac{4k_2 v_i^{\mathsf{tr}}(R - r_i^{\mathsf{tr}})}{\beta_r^2 \underbrace{(R - r_i^{\mathsf{tr}} - \tilde{x_i})^2}_{<(R - r_i^{\mathsf{tr}})^2} \underbrace{\left(k_1(R - r_i^{\mathsf{tr}}) + (2k_2 v_i^{\mathsf{tr}} - k_1)\tilde{x_i}\right)}_{<\left(\left(2v_i^{\mathsf{tr}}+1\right)k_2 - k_1\right)(R - r_i^{\mathsf{tr}})}}} \\ &> \underbrace{\frac{4k_2 - v_i^{\mathsf{tr}}}{\beta_r^2 \left(R - r_i^{\mathsf{tr}}\right)^2 \left(\left(2v_i^{\mathsf{tr}} + 1\right)k_2 - k_1\right)}}_{\diamondsuit} \end{aligned}$$

For any seller $i \in [n]$ and $k_1 > 0$, $\diamondsuit$ is a bijection: $k_2 \in [0, +\infty] \to [-\infty, +\infty]$. That is, $\forall i \in [n] \wedge k_1 > 0, \exists k_2 \in [0, +\infty] : \diamondsuit > 1$. This completes our verification for Assumption 1.