Essays In Quantitative Marketing

Vladimir Pavlov

University of Pennsylvania
Essays In Quantitative Marketing

Abstract
I study three questions related to competition and market design. In the first chapter, I study whether a peer-to-peer platform should set prices directly, or should it let sellers set prices while providing price recommendations. A platform can centralize prices and use exclusively available demand information, while price recommendations let sellers compete using their private information. On sharing economy platforms, for example, we observe a myriad of such pricing regimes. We investigate the implications of each pricing regime for the profits of platforms, buyers and sellers. When a platform recommends prices, it effectively plays the role of a sender in a multi-receiver cheap-talk game. In the second chapter, I study a particular instance of the trade-off between fairness and efficiency - a topic of increasing interest for marketers. After recreational cannabis legalization in 2012, Washington state policymakers had to choose whether to allocate retail licenses by lottery (a fair mechanism) or by auction (an efficient mechanism). They chose a lottery. Using transaction data from the Washington State Liquor and Cannabis Board, I estimate an equilibrium model of competition in the recreational cannabis market and use it to simulate the counterfactual auction allocation of licenses. I find that an auction would have increased total sales by 5% and reduced prices by 3%. As a result, the state lost $137M over ten years in tax revenue, which amounts to 0.39% of the state's annual budget. From the perspective of fairness, I find that under an auction, Black applicants are on average 21% less likely to receive a retail license and majority-White areas of the state reap disproportionately larger consumer benefits from the auction (20% increase vs 3% increase in consumer surplus for majority-Black and -Asian areas). In the third chapter, I document the effect vertical integration between broadcast television networks and major movie studios has on the studios' advertising strategies. I use a matching procedure together with a Difference-In-Differences estimator to separate the effect of vertical integration from the "content match". I find significant affiliation effect for ABC and Walt Disney, and FOX and 20th Century, but not NBC and Universal.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Marketing

First Advisor
Aviv Nevo

Subject Categories
Advertising and Promotion Management | Marketing

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/4137
ESSAYS IN QUANTITATIVE MARKETING

Vladimir Pavlov

A DISSERTATION

in

Marketing

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2021

Supervisor of Dissertation

______________________________
Aviv Nevo, George A. Weiss and Lydia Bravo Weiss University Professor

Graduate Group Chairperson

______________________________
Nancy R. Zhang, Ge Li and Ning Zhao Professor

Dissertation Committee

Eric T Bradlow, The K.P. Chao Professor

Ron Berman, Assistant Professor of Marketing
Dedicated to Michal
ACKNOWLEDGMENT

I would like to thank Aviv Nevo, Eric Bradlow and Ron Berman, the rest of the Whar-
ton Marketing faculty, as well as Julie Mortimer and Sylvia Hristakeva for their support
throughout my work on this dissertation. I thank Brett Hollenbeck for help with access to
WCLB data.
ABSTRACT

ESSAYS IN QUANTITATIVE MARKETING

Vladimir Pavlov

Aviv Nevo

I study three questions related to competition and market design.

In the first chapter, I study whether a peer-to-peer platform should set prices directly, or should it let sellers set prices while providing price recommendations. A platform can centralize prices and use exclusively available demand information, while price recommendations let sellers compete using their private information. On sharing economy platforms, for example, we observe a myriad of such pricing regimes. We investigate the implications of each pricing regime for the profits of platforms, buyers and sellers. When a platform recommends prices, it effectively plays the role of a sender in a multi-receiver cheap-talk game.

In the second chapter, I study a particular instance of the trade-off between fairness and efficiency - a topic of increasing interest for marketers. After recreational cannabis legalization in 2012, Washington state policymakers had to choose whether to allocate retail licenses by lottery (a fair mechanism) or by auction (an efficient mechanism). They chose a lottery. Using transaction data from the Washington State Liquor and Cannabis Board, I estimate an equilibrium model of competition in the recreational cannabis market and use it to simulate the counterfactual auction allocation of licenses. I find that an auction would have increased total sales by 5% and reduced prices by 3%. As a result, the state lost $137M over ten years in tax revenue, which amounts to 0.39% of the state’s annual budget. From the perspective of fairness, I find that under an auction, Black applicants are on average 21% less likely to receive a retail license and majority-White areas of the state reap disproportionately larger consumer benefits from the auction (20% increase vs
3% increase in consumer surplus for majority-Black and -Asian areas).

In the third chapter, I document the effect vertical integration between broadcast television networks and major movie studios has on the studios’ advertising strategies. I use a matching procedure together with a Difference-In-Differences estimator to separate the effect of vertical integration from the “content match”. I find significant affiliation effect for ABC and Walt Disney, and FOX and 20th Century, but not NBC and Universal.
TABLE OF CONTENTS

ACKNOWLEDGMENT .................................................. iii
ABSTRACT ........................................................... iv
LIST OF TABLES ...................................................... viii
LIST OF ILLUSTRATIONS ................................ .......... x

CHAPTER 1 : Price Manipulation in P2P Markets and Sharing Economy . . . . 1
  1.1 Introduction .................................................... 1
  1.2 Contribution to Literature .................................... 6
  1.3 Model and Pricing Benchmark ................................. 7
  1.4 Price Manipulation with Cheap Talk ......................... 16
  1.5 Market Implications ........................................... 23
  1.6 Impact of Search Technology $\alpha$ ........................ 27
  1.7 Conclusion .................................................... 29

CHAPTER 2 : The Price of Fairness: Equity and Efficiency in Retail License Allo-
cation ............................................................. 32
  2.1 Introduction .................................................... 32
  2.2 Background and Data ......................................... 40
  2.3 Model .......................................................... 45
  2.4 Estimation ...................................................... 52
  2.5 Estimation Results ............................................. 56
  2.6 Counterfactual: Using Auctions Instead of Lotteries ........... 60
  2.7 Conclusion .................................................... 72
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>Summary statistics for retailers by the owner’s predicted racial identity.</td>
<td>45</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>First stage results, standard errors in parentheses</td>
<td>56</td>
</tr>
<tr>
<td>TABLE 3</td>
<td>Demand estimation results, standard errors in parentheses</td>
<td>57</td>
</tr>
<tr>
<td>TABLE 4</td>
<td>Summary statistics for the prediction error in quantity sold</td>
<td>59</td>
</tr>
<tr>
<td>TABLE 5</td>
<td>Market outcomes under lottery and auction</td>
<td>62</td>
</tr>
<tr>
<td>TABLE 6</td>
<td>Retailer selection under lottery and auction</td>
<td>64</td>
</tr>
<tr>
<td>TABLE 7</td>
<td>Predicted race of license owners under auction and lottery</td>
<td>65</td>
</tr>
<tr>
<td>TABLE 8</td>
<td>Consumer benefits of auction on the block level, divided by the race of majority of residents.</td>
<td>66</td>
</tr>
<tr>
<td>TABLE 9</td>
<td>Consumer benefits of auction on the block level, divided by decile of share of white residents</td>
<td>67</td>
</tr>
<tr>
<td>TABLE 10</td>
<td>Major television and movie studio ownership structure in 2011-2013</td>
<td>87</td>
</tr>
<tr>
<td>TABLE 11</td>
<td>Number of films in each genre by parent company</td>
<td>88</td>
</tr>
<tr>
<td>TABLE 12</td>
<td>Major studios’ TV advertising allocation, affiliated studio-network pairs in bold.</td>
<td>89</td>
</tr>
<tr>
<td>TABLE 13</td>
<td>Top 20 cable networks used for matching</td>
<td>92</td>
</tr>
<tr>
<td>TABLE 14</td>
<td>Top 20 categories of products matched to movies by parent company</td>
<td>93</td>
</tr>
<tr>
<td>TABLE 15</td>
<td>Top 20 categories of matched products by movie genre</td>
<td>94</td>
</tr>
<tr>
<td>TABLE 16</td>
<td>Regression results for Comcast movies and matched products.</td>
<td>95</td>
</tr>
<tr>
<td>TABLE 17</td>
<td>Regression results for Disney movies and matched products.</td>
<td>96</td>
</tr>
<tr>
<td>TABLE 18</td>
<td>Regression results for News Corp movies and matched products.</td>
<td>96</td>
</tr>
<tr>
<td>TABLE 19</td>
<td>Falsification exercise: estimated effects from treating Warner Bros as an affiliated studio.</td>
<td>97</td>
</tr>
<tr>
<td>TABLE 20</td>
<td>Falsification exercise: estimated effects from treating Paramount as an affiliated studio.</td>
<td>97</td>
</tr>
</tbody>
</table>
TABLE 21: Falsification exercise: estimated effects from treating Columbia as an affiliated studio. ................................. 97
TABLE 22: Regression results for Comcast movies and matched products, heterogeneous effect for 10% most heavily advertising films. .... .... 98
TABLE 23: Regression results for Walt Disney movies and matched products, heterogeneous effect for 10% most heavily advertising films. .... .... 98
TABLE 24: Regression results for News Corp movies and matched products, heterogeneous effect for 10% most heavily advertising films. .... .... 99
TABLE 25: Regression results for Comcast films, with treatment heterogeneity for top two movie genres produced by the studio: comedy and action. 99
TABLE 26: Regression results for Disney films, with treatment heterogeneity for animation and family movies. ................................. 100
TABLE 27: Regression results for News Corp films, with treatment heterogeneity for comedy and drama movies - top two genres produced by the studio. 100
TABLE 28: Summary statistics for the prediction error in prices .................. 115
TABLE 29: Regression results for Comcast movies, using only other films as the set of possible matches. ................................. 119
TABLE 30: Regression results for Walt Disney movies, using only other films as the set of possible matches. ................................. 120
TABLE 31: Regression results for News Corp movies, using only other films as the set of possible matches. ................................. 120
LIST OF ILLUSTRATIONS

FIGURE 1: Pricing regimes that the platform and the sellers prefer depending on the values of $q$ and $\bar{v}$. Other parameters: $\alpha = 0.45$. 24

FIGURE 2: Pricing regimes that the buyers prefer depending on the values of $q$ and $\bar{v}$ when $\alpha = 0.45$ and 27

FIGURE 3: Number of possible cheap-talk intervals ($n^*$) as a function of $\alpha$ and $\bar{v}$ when 28

FIGURE 4: Historical growth of cannabis retail sales in Washington 42

FIGURE 5: Time series of cannabis retail prices in Washington 42

FIGURE 6: Time series of cannabis retail margins in Washington 43

FIGURE 7: Histogram of average retailer margins 44

FIGURE 8: Distribution of retailer types 58

FIGURE 9: Percentage increase in consumer surplus from the auction as a function of median household income decile. 68

FIGURE 10: Probability of at least one store being present in an average block group as a function of median household income decile. 68

FIGURE 11: Probability of winning conditional on location, from the least to the most successful location, for auction and lottery 69

FIGURE 12: Probability of winning conditional on type, from the least to the most successful type, for auction and lottery 69

FIGURE 13: Areas that are better off under auction (yellow) vs under lottery (purple) 71

FIGURE 14: Inequality in average consumer surplus, by block group, under auction and lottery 71
FIGURE 15: Market share of Disney as a function of level of integration $\lambda$, $v = 1$, 
$\bar{u}_D = 2, \bar{u} = 3/2$. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 82

FIGURE 16: Welfare as a function of level of integration $\lambda$, $v = 1$, $\bar{u}_D = 2, \bar{u} = 3/2$. 85

FIGURE 17: Distribution of prediction errors for individual retailers . . . . . . 116
FIGURE 18: Distribution of prediction errors on market level . . . . . . . . . . 117
FIGURE 19: Market-level prediction errors vs number of firms in the market . 117
FIGURE 20: Error in quantity for the entrant vs the number of firms in the market 118
CHAPTER 1 : Price Manipulation in P2P Markets and Sharing Economy ¹

1.1. Introduction

Platforms that operate peer-to-peer (P2P) markets can influence the interaction between sellers and buyers through their platform design. Among many examples, Lyft and Uber control pricing to facilitate rides between riders (buyers) and drivers (sellers), Airbnb controls search results and recommends pricing to influence matching between hosts (sellers) who rent their houses to guests (buyers), and LendingClub assigns a credit worthiness score to borrowers (sellers) who are asking for a loan from investors (buyers).

Since the P2P structure is prevalent in many industries, it is not surprising that there is no “one size fits all” market design of a P2P market. Platforms differ in many aspects including how search results are presented to buyers, their fee structure, how much choice buyers and sellers have and whether sellers can promote their offerings for an additional fee. Many of these differences arise from the choices platforms make when using information about consumer demand and seller competition to maximize their profits (if that is what they are doing).

In many P2P markets, sellers often find it difficult to pick prices for their products because of uncertainty about demand and competition. Equilibrium price levels, however, have a dramatic impact on the profits of the platform. Higher prices will lead to less transactions but with a higher margin, while low prices may increase the number of purchases but erode profit. For example, when Airbnb initially launched their platform, sellers were setting high prices that lowered the number of transactions, user satisfaction and platform revenue. Consequently Airbnb introduced a price recommendation tool for hosts in 2013 which they later improved in 2015 (Hill, 2015). Because sellers set their prices based on beliefs they have about buyer demand, the platform can influence competition and price levels through supplying information to sellers or through coordinating prices directly.

¹Based on work with Ron Berman
A second factor that impacts the long-term profitability of the platform is the quality of matches achieved in realized transactions. Low-quality matches will lead to long-term churn of buyers who switch to competing options. The platform can use its information about buyer preferences and seller differentiation to influence the quality of matches offered to buyers, alongside the prices achieved in these transactions. In this paper we analyze how the information design of the platform can maximize its profits through manipulating the realized price levels, levels of competition and match quality.

Researchers recently devoted substantial attention to analyzing selling mechanisms on P2P platforms, with particular focus on using auctions vs. posted prices (Hammond, 2010, 2013; Bauner, 2015; Einav et al., 2015, 2018; Waisman, 2017). Another stream of research looks at dynamic pricing as a tool to influence the supply of sellers in the market (Guda and Subramanian, 2019).

Less attention was given to which party sets the prices on the platform. While Lyft’s algorithm sets the price for each ride (centralized pricing), Airbnb hosts and eBay sellers are free to set their own prices (competitive pricing). However, even in competitive markets, platforms sometimes participate by providing a price recommendation to sellers/hosts. Following its introduction of “price tips” in 2015, for example, Airbnb developed the recommendation tools further by introducing “smart pricing” in 2017, which lets the host set the maximum and the minimum price of a stay and the platform adjusts the price of the listing in response to predicted changes in demand.

When choosing whether to centralize pricing or let sellers compete, the platform has to consider two key factors. First is the amount of demand information the platform possesses relative to the information sellers have. In the case of Airbnb, for example, how good is the pricing algorithm at predicting demand compared to the hosts themselves? If the platform decides to centralize pricing, the prices will not reflect the private information that sellers have. But if the platform chooses to let sellers compete without providing them with information, the sellers cannot use demand-related data available to the platform to
assist in their pricing decisions. Recommending prices may alleviate this trade-off partially; it lets the platform share some of its information with the sellers, while giving the sellers flexibility to compete. As we will show, these choices may alter the level of competition in the market, influencing the equilibrium price levels and match quality. The second factor the platform needs to consider is that in a competitive market sellers will set relatively low prices, while centralized pricing allows the platform to extract more of the consumer surplus. A possible solution to this tradeoff is to let sellers compete while providing them with price recommendations that align with the platform's goals. However, as we will show, price recommendations constitute cheap talk by the platform, posing a challenge to the usefulness of this strategy if the recommendations are not credible and ignored by the sellers.

Our goal in this paper is to describe when a platform would prefer to centralize or decentralize pricing and the implications of this decision for the platform, sellers and buyers. Our results can help platform designers make informed decisions regarding pricing regimes and information design. Because we also consider the effects of these choices on consumer surplus, our results can also guide policymakers and regulators considering the regulation and impact of P2P markets.

To study these questions, we construct a theoretical model of segmented competition between two differentiated sellers who sell imperfect substitutes on a platform to buyers. Buyers choose whether to buy a product in the market or pick an outside option. The platform can design a search technology that limits the buyers ability to buying from only one seller, or two. The platform also chooses whether to set prices for the sellers or to let the sellers compete with or without a price recommendation. If the sellers are allowed to compete, they can set prices for their product while taking the recommendation of the platform into account.

There are two sources of uncertainty in our model. First, each seller has private information about the quality of her product, which affects the utility buyers receive from the product.
Second, there is an aggregate market-level shock to willingness to pay (e.g., how many visitors to a certain city are budget-conscious tourists and how many are business travelers with expense accounts). This shock is observed by the platform, but not by the sellers.

Like most real-life platforms, we assume that the platform receives a fixed-percentage fee of the sellers’ revenues. This means that while the platform wants to maximize the joint revenue of the sellers, each seller seeks to maximize their own profit and does not internalize the sales they take away from their competitor by lowering their price. Because of this misalignment of incentives between the sellers and the platform, the sellers will not follow a price recommendation by the platform blindly. The sellers will form rational expectations of the platform’s strategy, i.e., in which state of the world the platform will choose to recommend each price. This messaging game where the platform recommends prices and the sellers have misaligned incentives is an instance of classic cheap talk (Crawford and Sobel, 1982). However, unlike Crawford and Sobel (1982) and most of the cheap talk literature, our model has multiple receivers (sellers) that interact with each other (i.e., compete) and the outcome of this interaction determines the payoff of the sender. In a standard cheap talk game, the misalignment in incentives between the sender and the receiver is exogenous, while in our model it is endogenous and stems from the difference in market power between the platform and the sellers, as well as the level of competition among the sellers. Both of these are influenced by the price recommendations of the platform.

We find that the platform should choose to centralize pricing if there is little uncertainty about the quality of sellers’ products. On the other hand, if this uncertainty is large and the variance of the aggregate demand shock observed by the platform is large, the platform should recommend a price. If the variance of the aggregate shock is small, price recommendations cannot be credible in equilibrium and the platform should let the sellers set their own prices. The intuition is that the agents that posses the more valuable demand information should set the prices that reflect that information.

To provide an example of such a setting, consider an Airbnb market in a particular city. The
willingness to pay of buyers depends on how many business travelers are looking to book in this market, which the platform can observe. The private information each seller (host) has is the (unobserved by the platform) quality of apartment offered. For example the quality of the view or location is often hard for platforms to assess. If both sources of uncertainty are small, e.g., it is either a place where people only come for business that has no good views at all or a place where all views are great and that attracts primarily tourists, the platform should centralize pricing to leverage its position as a monopolist. If both sources of uncertainty are large, e.g., there are different kinds of travelers in this market and some apartments can have extremely nice views while others are facing a brick wall, the platform should recommend a price and allow sellers to use both sources of information for pricing.

When we focus on the benefit of pricing regimes for sellers, we find a few surprising results. High quality sellers surprisingly exhibit a stronger preference for centralized pricing, while low quality sellers have a stronger preference for recommendation and competition. These differences stem from how competition affects the levels of demand and prices when a high quality seller faces a low quality seller.

We also find that buyer surplus is almost always maximized under competitive pricing by the sellers. Unlike sellers, buyers do not benefit when prices adjust with the state of the demand. If the price is too high, buyers can always take an outside option with a limited downside. But if the price is low when demand is high, they will enjoy a large surplus. Therefore, buyers are better off when prices respond the least to changing demand. In most cases, this happens under competitive pricing, as the sellers only take into account the information about their own quality. Price recommendations always hurt buyers compared to the fully competitive case, as it increases price variance without lowering the average price. Under centralized pricing, the average price is higher, but in extreme cases when the variance of the aggregate shock is small (all visitors are tourists), but the range of possible qualities is large (the views are superb or terrible), then the buyers prefer centralized pricing.

One emerging result from our analysis shows that platforms are generally better off (in
terms of profit) with centralized pricing and search technology that gives customers many options to choose from. However, if platforms also want to take their growth into account, price recommendations provide a better avenue to achieve both growth and profit, because centralized pricing excludes consumers from the market. A second emerging result is that unlike much of the previous research, we find that centralizing prices through a platform is not always profit maximizing for both the platforms and sellers. Moreover, there are cases where price recommendations from a platform may not be credible, and the platform might be better off not offering them at all.

Following a review of the literature, we present our game theoretic model. We then follow with an analysis of a benchmark case with a platform that can choose between centralized pricing and competition. This analysis serves as a stepping stone to analyze the cheap-talk game where platforms recommend prices. We conclude with an analysis of the market implications for different stakeholders. All proofs are relegated to Appendix A.1.

1.2. Contribution to Literature

Our work contributes to three streams of literature. First, our paper adds to the growing literature on P2P platforms (see, e.g., Narasimhan et al. (2018) and Eckhardt et al. (2019) for surveys on related research in marketing). The substantial research on platform design and impact (e.g., Einav et al., 2016; Horton and Zeckhauser, 2016; Jiang and Tian, 2016; Zervas et al., 2017; Ke et al., 2017; Fradkin, 2017; Fradkin et al., 2018; Guda and Subramanian, 2019) has focused on measuring the impact of collaborative consumption on the market, as well as explored how different features of platform design affect market outcomes. Within this literature, the research on platform pricing has mostly considered auctions vs. posted prices, or the impact of dynamic pricing on matching buyers and sellers. Auctions, however, are not a natural choice for many platforms, while dynamic pricing requires exerting pricing controls and having dynamic information that many platforms might not possess. For this reason we focus on two common general mechanisms - centralized pricing where the platform sets pricing for all sellers vs. competitive pricing where the platform allows
sellers to set their own prices, with or without a price recommendation.

Second, we add to the theoretical work on strategic communication (Milgrom, 1981, 2008; Crawford and Sobel, 1982; Sobel, 2013), which has recently been applied in marketing contexts on persuasive communication (Gardete, 2013; Chakraborty and Harbaugh, 2014). Interestingly, cheap talk has rarely been applied to the analysis of a market with many competing receivers. The work of Kim and Kircher (2015), for example, has many senders who send cheap-talk messages, while our work looks at a sender trying to coordinate a market using cheap talk. We prove that the results of Crawford and Sobel (1982) are robust to introducing competition in our model: we find that all possible equilibria have “coarse” communication, i.e., the platform recommends a range of prices instead of a single price and the platform (the sender) and the sellers (receivers) benefit when more fine-grained communication is possible.

Finally, our paper is related to the literature on oligopolistic competition under uncertainty (Klemperer and Meyer, 1986, 1989; Gardete, 2016). While in these works the level of competition is determined by exogenous uncertainty competitors face, our research extends these works to a scenario in which a market designer can control the level of uncertainty facing competitors and thus influence the level of the competition to her benefit using cheap talk.

1.3. Model and Pricing Benchmark

There are three types of players who interact in a one-shot game: buyers, sellers and a platform. The mass of $2M$ buyers are distributed uniformly on the real line in the range $[-M, M]$ where $M$ is large. The buyers visit a P2P platform to buy a product from one of two potential sellers. Seller 1 is positioned at $-1$ and seller 2 is positioned at 1.\footnote{The symmetry around 0 simplifies the exposition vs. a standard Hotelling model with locations 0 and 1.} A buyer located at $x \in [-M, M]$ has demand for up to one product. If they choose to buy a product
from seller $i$, their utility is:

$$u_i(x) = v + q_i - p_i - d_i(x) \quad (1.1)$$

Where $p_i$ is the price of product $i$, $d_i(x)$ is the distance between the seller and buyer and $q_i$ is the quality of the product which equals $-q$ or $q$ with $q > 0$. Because sellers 1 and 2 are located at 1 and $-1$ respectively, the distances equal $d_1(x) = |x + 1|$ and $d_2(x) = |x - 1|$.

The buyer’s willingness to pay $v$ is drawn from an a-priori uniform distribution $U[v, \bar{v}]$ with $v > 0$. We assume that the platform has more information about the realization of $v$ than the sellers. If buyers are business travelers, for example, their willingness to pay might be higher than budget conscious tourists. A platform will be able to observe if searches for listings in a specific city, for example, come mostly from business travelers. Buyers will also know their realization of $v$ before buying the product. Sellers, in contrast, are not exposed to the search process on the platform, and hence will have less information about $v$. For simplicity, we assume that the platform observes the realization of $v$ and the sellers do not.

The value of $q_i$ is private information of each seller. It is drawn independently with equal a-priori probability of being $q$ or $-q$. We assume that buyers can observe $q_i$ before purchasing the product, but the platform cannot. Although we abstract away from the details of such a setup, this is consistent with a signaling game where buyers can learn the quality from observed prices before making their purchase decisions, but the platform cannot learn these qualities before making its own decisions. Appendix A.2 provides a detailed signaling model in which prices signal quality in equilibrium under competition, but not under centralized pricing.

Finally, we assume that the outside option is the same for all buyers and is normalized to 0. The outside option can capture the utility from going to a competitor (taxi, public transit or Uber instead of Lyft) or from not making any purchase.

Before visiting the platform, buyers are not aware of the sellers and hence cannot buy from
them. Once consumers visit the platform, buyers use search tools to find their preferred product. We assume a simple search technology: a share $\alpha < 1/2$ of buyers discover seller 1 only, another group of size $\alpha$ discover seller 2 only, and the remaining mass of $1 - 2\alpha$ buyers discovers both sellers. We call the first two groups “captives” as they can only buy from one seller or pick the outside option. We call the buyers who are aware of both sellers “comparison shoppers” (or shoppers). Once consumers visit the platform, they see the prices, distances and qualities of each seller they are aware of. The consumers pick the option that gives them the highest utility among the sellers or the outside option. Initially, we assume that $\alpha$ is exogenous. Later, we compare the choices of different values of $\alpha$ in Section 1.6.

The platform receives a fixed share $\phi$ of the sellers’ revenues that is set exogenously. This revenue-sharing arrangement mimics many of the contracts in the P2P universe. Hence seller $i$’s expected profit is:

$$\pi_i = \mathbb{E}_v[(1 - \phi)p_iD_i(p_1, p_2)]$$  \hspace{1cm} (1.2)

while the platform’s expected payoff is:

$$\pi_P = \mathbb{E}_{q_1, q_2}[\phi(p_1D_1(p_1, p_2) + p_2D_2(p_1, p_2))]$$  \hspace{1cm} (1.3)

where $D_i(p_1, p_2)$ is the realized demand, which is equal the total mass of consumers buying from seller $i$.

From the profit functions it is evident that the sellers and the platform have misaligned incentives. The platform would like to maximize the joint revenue of the sellers, while each seller would like to maximize its own profit. Moreover, the competition might affect the pricing incentives of the sellers, while the platform does not care which seller sells the product as long as a transaction is made. Hence the incentives of the platform and the sellers are not perfectly aligned with respect to setting prices.
A second factor that affects the profit of the sellers and the platform is information asymmetry. If the platform sets the prices, it cannot use the information that the sellers have about $q_i$. If the sellers set prices, they are uncertain about the willingness to pay $v$ unless they receive information from the platform.

Our main goal is to analyze the usage of price recommendations by the platform to profitably manipulate the equilibrium profits, price levels and matching quality in the market. To achieve this goal, we initially analyze two benchmark pricing strategies to gain insights about the model and serve as a stepping stone for the analysis of the price recommendation strategy. We will analyze the equilibria of the following three cases in terms of platform and seller profits, consumer surplus, and quality of matching:

1. Centralized Pricing (CP) - The platform sets prices for both sellers.

2. Competition (C) - The platform lets the sellers set their own prices without providing them information.

3. Recommendation (R) - The platform recommends a non-binding price to the sellers, and each seller sets their own price.

As we discuss in detail later, any recommendation message the platform sends to sellers is effectively a signal about the value of $v$, because $v$ is the only payoff-relevant private information that the platform has. In other words, all price recommendations that the platform can provide in our model are functions of $v$ and are therefore isomorphic to a “direct” message communicating the value of $v$.

The three cases above cover the full range of actions the platform can take to influence the sellers in our model. To finalize the model, the timing of the game is as follows:

1. The platform selects a pricing and recommendation strategy.

2. Nature draws $v$ (observed by the platform and buyers), $q_1$ (observed by seller 1) and $q_2$ (observed by seller 2).
3. The platform gives a price recommendation to the sellers if it decided to do so.

4. Prices are set by the platform or by the sellers.

5. Buyers visit the platform. They learn $d_i$, $p_i$ and $q_i$ for one or two sellers.

6. Buyers make their purchase decisions and payoffs are realized.

At step 3, if the platform elects to not recommend a price, we can assume that it is giving an uninformative recommendation (e.g., recommends a random price independent of the state of the world $v$ or the same price in every state of the world).

1.3.1. Centralized Pricing vs. Competition

To analyze the benchmark cases of centralized pricing and competitive pricing, we first derive the demand experienced by each seller when the prices are $p_1$ and $p_2$.

For captive consumers, each buyer chooses between seller $i$ and the outside option. There are two buyers who are indifferent between buying and not buying, equidistant to the left and to the right of $-1$ (seller 1) or 1 (seller 2). The demand from captives is then the mass of buyers between these two points:

$$D_i^{cap}(p_i) = 2(v + q_i - p_i) \quad (1.4)$$

To find the demand from comparison shoppers, we first make an assumption that facilitates tractability of the analysis:

**Assumption A1.** (i) $v > 2q + 3 - 2\alpha$ and (ii) $q < 1$.

The first part of Assumption A1 is a standard full coverage assumption for shoppers in the $[-1, 1]$ interval, and makes sure these shoppers buy from either firm 1 or 2. The second part implies that the difference in quality between sellers is not so high that comparison shoppers always buy only from one seller. Effectively, it guarantees that if the prices set by
the sellers are equal, each seller will receive some demand from comparison shoppers even if
qualities are different. Relaxing these assumptions will not change the results qualitatively,
but will make the analysis less tractable.

Using assumption A1, the demand of shoppers is:

\[ D_i^s(p_i, p_{-i}) = \frac{q_i - q_{-i} - p_i + p_{-i}}{2} + v + 1 + q_i - p_i \]  (1.5)

Combining the demands from both segments the total demand for seller \( i \) can be written
to yield a linear and differentiated Bertrand model (Klemperer and Meyer, 1986, 1989):

\[ D_i(p_i, p_{-i}) = \frac{2v + (3 - 2\alpha)q_i - (1 - 2\alpha)q_{-i} + 2(1 - 2\alpha)}{2} - \frac{3 - 2\alpha}{2}p_i + \frac{1 - 2\alpha}{2}p_{-i} \]  (1.6)

The differentiation in the model stems from the difference in price sensitivities of consumers
buying from a specific firm. Because each firm’s own price influences also the captive
segment, the demand each firm sees is more elastic with respect to changes in its own price
compared to changes in the competitor’s price.

Solving for the profit-maximizing price results in the following:

**Proposition 1.** When using centralized pricing:

- The unique profit-maximizing price is: \( p_1^* = p_2^* = p^*(v) = \frac{v+1-2\alpha}{2} \).

The optimal centralized price increases with \( v \), but decreases with \( \alpha \).

- The maximum centralized profit is: \( \pi^{CP}_P = \phi \frac{(v+1-2\alpha)^2}{2} \).

The optimal centralized profit increases in \( v \) and decreases with \( \alpha \).

- The maximum expected centralized profit is:

\[ \mathbb{E}(\pi^{CP}_P) = \phi \frac{(\bar{v} + 1 - 2\alpha)^3 - (v + 1 - 2\alpha)^3}{6(\bar{v} - v)} \]
The expected centralized profit increases in $\bar{v}$ and $v$, and decreases with $\alpha$.

- The ex ante expected profit of high and low type sellers is:

$$
\mathbb{E}(\pi_H^{CP}) = (1 - \phi) \left( \mathbb{E}(\pi_P^{CP}) \left( \frac{q(3 - 2\alpha)(\bar{v} + v + 2(1 - 2\alpha))}{8} \right) \right) \\
\mathbb{E}(\pi_L^{CP}) = (1 - \phi) \left( \mathbb{E}(\pi_P^{CP}) \left( \frac{q(3 - 2\alpha)(\bar{v} + v + 2(1 - 2\alpha))}{8} \right) \right)
$$

Proposition 1 shows that prices and profits increase, as expected, with the average willingness to pay $v$. A surprising result is that the profit decreases with $\alpha$, as intuition might suggest that the platform could gain the most by exposing each buyer to only one product, and use monopoly pricing for each product. This intuition breaks when the platform has uncertainty over $q_i$. When the sellers differ in quality, if most buyers are captives (high $\alpha$), those aware of the low quality seller will only buy if the price is low. The platform, however, is constrained to setting the same price for both products. If these buyers are made aware about the other seller, however, they may be willing to make a purchase at a higher price from a high quality seller. Hence, it is in the interest of the platform to make more buyers informed and decrease $\alpha$.

Turning the attention of analysis to competition, seller $i$ with type $\tau$ will set a price $p_{i\tau}$ to maximize their expected profit. We look for a symmetric subgame perfect equilibrium, and hence can denote the equilibrium strategy of sellers as $p_{C\tau}$ for type $\tau \in \{H, L\}$ and drop the subscript $i$. The profit of a seller of type $\tau$ who sets a price $p_\tau$ (not necessarily equal to $p_{C\tau}$) is:

$$
\mathbb{E}(\pi_\tau) = \frac{1}{2} (1 - \phi) p_\tau \mathbb{E}_v(D_v(p_\tau, p_{C_H}) + D_v(p_\tau, p_{C_L}))
$$

(1.7)

where $D_v$ denotes the demand of a seller with type $\tau$.

Solving for the equilibrium results in the following:

**Proposition 2.** Under price competition between the sellers:
• Equilibrium prices are:

\[
p_C^H = \frac{q}{2} + \frac{2E(v) + 2(1 - 2\alpha)}{5 - 2\alpha} \tag{1.8}
\]
\[
p_C^L = \frac{-q}{2} + \frac{2E(v) + 2(1 - 2\alpha)}{5 - 2\alpha} \tag{1.9}
\]

The price \( p_C^H \) increases with \( q \) and \( p_C^L \) decreases with \( q \). Both prices increase in \( \alpha \) if \( \bar{v} + \bar{v} > 8 \) and decrease in \( \alpha \) otherwise.

• The platform’s profits are:

\[
\pi_P = \phi \frac{q^2(3 - 2\alpha)}{4} - \phi \frac{16E(v)}{4(5 - 2\alpha)^2} + \frac{16E(v)}{4(5 - 2\alpha)^2} \left( 2E(v) - (5 - 2\alpha)v - 3 + 8\alpha - 4\alpha^2 \right)
\]

• The platform’s ex ante profits are:

\[
E(\pi_P) = \frac{\phi(3 - 2\alpha)(4(\bar{v} + \bar{v} + 2(1 - 2\alpha))^2 + q^2(5 - 2\alpha)^2)}{4(5 - 2\alpha)^2}
\]

• The sellers’ ex ante profits are:

\[
E(\pi_H) = \frac{(1 - \phi)(3 - 2\alpha)(2(\bar{v} + \bar{v} + 2(1 - 2\alpha)) + (5 - 2\alpha)q)^2}{8(5 - 2\alpha)^2}
\]
\[
E(\pi_L) = \frac{(1 - \phi)(3 - 2\alpha)(2(\bar{v} + \bar{v} + 2(1 - 2\alpha)) - (5 - 2\alpha)q)^2}{8(5 - 2\alpha)^2}
\]

Proposition 2 shows that prices are linear in the beliefs of the sellers about the expectation of \( v \). When we analyze price recommendation using cheap talk in the next section, this feature will come into play as the platform will want to influence the resulting equilibrium prices through influencing the beliefs of sellers. Specifically, the platform’s profit is quadratic in the beliefs of the sellers, first increasing and then decreasing.

When we compare the equilibrium prices and profits between centralized pricing and com-
petition, we find the following:

**Corollary 1.**

- The prices $p_L^C$ and $p_H^C$ are higher than the centralized price when $\bar{v}$ is high enough and $v$ is low enough.

- The profits of the platform are higher than under centralized pricing when

$$q > \sqrt{\frac{2}{3 - 2\alpha}} \left( v + \frac{(1 - 2\alpha)^2 - 2(\bar{v} + v)}{5 - 2\alpha} \right)$$

- There is a $\tilde{q}$ such that the platform makes a higher ex-ante profit without a price recommendation compared to centralized pricing if and only if $q > \tilde{q}$.

- There is a $\tilde{q}_H$ ($\tilde{q}_L$) such that a high (low) quality seller makes a higher ex-ante profit without a price recommendation compared to centralized pricing if and only if $q > \tilde{q}_H$ ($q > \tilde{q}_L$).

Corollary 1 has two interesting findings. First, prices under competition may be higher than prices set by a centralized planner. When $\bar{v}$ is high, and the realization of $v$ is low, sellers will set too high prices because they expect higher demand than what is realized, and as a result will lower the platform’s profit. If the platform could affect the beliefs of sellers about $v$, it might be able to better influence this competition to its benefit.

The second interesting finding is that in scenarios where $q$ is high, and there is substantial uncertainty about the difference in seller quality, it is more beneficial for the platform to let sellers compete than to set prices for them. This is in contrast to the previous literature that found that coordinating centralized prices is beneficial for a platform as it softens competition among sellers and increases sale prices. In a world where there is sufficient uncertainty about the quality of sellers on a platform, the platform should relinquish the pricing power to the players that hold the most uncertain information.
Having established the considerations for picking among the two basic pricing models, in the next section we turn to analyze price recommendations. These recommendations allow the platform to influence the beliefs of sellers about \( v \), and thus manipulate the benefits of competition for the platform.

1.4. Price Manipulation with Cheap Talk

When sellers set their own prices they integrate over their beliefs about \( v \) to maximize their expected profits (Equation (1.6)). The platform can try to influence the sellers’ decision by providing them with a price recommendation that the sellers will incorporate into their decisions. Providing a price recommendation and providing information about the value of \( v \) are equivalent because setting prices is the only action sellers can take, and \( v \) is the only missing piece of information sellers need from the platform. If the platform chooses to provide (possibly inaccurate) information about \( v \), the sellers can back-out the real value of \( v \) consistent with an equilibrium strategy of the platform, and make a pricing decision. Similarly, if the platform provides a price recommendation (and not a direct message about the value of \( v \)), the sellers will infer the values of \( v \) which are consistent (in equilibrium) with the platform’s recommendation.

We therefore assume that the platform’s strategy is a (possibly non-deterministic) mapping from the interval of possible realizations of \( v \in [v, \overline{v}] \) to a message space on the same interval. In other words, the platform observes the realization of \( v \) and reports to the sellers some plausible value \( m(v) \in [v, \overline{v}] \), which may or may not coincide with the actual realization.

An important feature of the model is that the message \( m(v) \) is costless for the platform (the Sender) to send, and that the platform’s incentives are misaligned with the sellers (the Receivers). Sellers have an incentive to lower prices to respond to competition and maximize their own profits, while the platform would like sellers to maximize their joint profit, which often means increasing their prices from a competitive level. This is an instance of a cheap-talk game (Crawford and Sobel, 1982), but unlike the extant cheap-talk literature, our
model features multiple receivers who interact strategically with each other. Our analysis also tries to answer whether cheap talk can be both a credible and a profitable equilibrium strategy with competing receivers.

A second interesting insight is that a “babbling equilibrium”, which always exists in cheap talk games, coincides in our model with the competition scenario we analyzed in the previous section. In such an equilibrium, the message sent by the platform is uninformative, i.e., it is statistically independent from the realization of $v$. Examples of such strategies would be to always recommend the same price, or to report a random value of $v$ to sellers. The sellers will then ignore the message and rely on their prior beliefs over $v$ when setting prices.

To understand what actions the platform should take, we first analyze the response of sellers to a message $m$ in the pricing subgame. When receiving a recommendation $m$, sellers will update their beliefs (using Bayesian updating) about the distribution of $v$. When updating their beliefs sellers will take into account the equilibrium strategy $m(v)$ used by the platform to narrow the possible values of $v$ to those consistent with the message $m$. The resulting equilibrium prices depend only on the updated expected value of $v$, $E(v|m)$ as shown in the following Lemma:

**Lemma 1.** Given a platform’s messaging strategy $m(v)$ and after receiving a message $m$, the unique equilibrium prices that sellers set are:

$$
 p_R^H = \frac{q}{2} + \frac{2E(v|m) + 2(1-2\alpha)}{5-2\alpha} \tag{1.10}
$$

$$
 p_L^R = -\frac{q}{2} + \frac{2E(v|m) + 2(1-2\alpha)}{5-2\alpha} \tag{1.11}
$$

where $E(v|m) = \frac{\int_{v:\ m=m} v \, dv}{\int_{v:\ m=m} dv}$.

When we compare to the results of proposition 2, it is notable that the recommendation of the platform affects the prices through the expectation linearly, and that if both types of sellers believe the expected value of $v$ is higher, they will set higher prices.
Because the platform influences the decision of the sellers by communicating a value for $v$, we can calculate the profit of the platform as a function of the true state $v$ and the seller’s expectations induced by $m$:

$$
\pi^R_{P}(v,m) = \frac{q^2(3-2\alpha)}{4} - \frac{16(\mathbb{E}(v|m) + 1 - 2\alpha)(2\mathbb{E}(v|m) - (5 - 2\alpha)v - 3 + 8\alpha - 4\alpha^2))}{4(5-2\alpha)^2}
$$

$\pi^R(v,m)$ is quadratic in $\mathbb{E}(v|m)$ and linear in $v$. Consequently every state $v$ has a value $\mathbb{E}^*(v)$ that maximizes the payoff of the platform:

$$
\mathbb{E}^*(v) = \frac{v(5 - 2\alpha) + (1 - 2\alpha)^2}{4}
$$

This expectation does not equal to $v$ itself and is in fact always larger than $v$, hence the platform would like to inflate the sellers’ expectations of $v$ through the recommendation. However, as sellers are rational and anticipate this strategy of the platform, that is impossible.

Given this limitation, we show in the next Lemma (based on Lemma 1 of Crawford and Sobel (1982)) that only a finite set of of beliefs can be induced in equilibrium, which implies that the true value of $v$ cannot be communicated, and only an indication of ranges of values of $v$ can be sent as a message:

**Lemma 2.** If for every message $m$ the values $v \neq \mathbb{E}^*(v|m)$, then there exists an $\varepsilon > 0$, such that for any two equilibrium messages $m_1$ and $m_2$ that induce different beliefs $\mathbb{E}(v|m_1)$ and $\mathbb{E}(v|m_2)$, the difference is at least $\varepsilon$, i.e., $|\mathbb{E}(v|m_1) - \mathbb{E}(v|m_2)| > \varepsilon$. Moreover, the set of expectations that can be induced in equilibrium is finite.

Lemma 2 shows that whenever two messages induce different equilibrium beliefs, those beliefs will have at least some minimal distance between them. In order words, the platform cannot induce a continuous set of beliefs and will have “jumps” between them. The intuition behind this result is that because the platform’s incentives and the seller incentives differ, the platform will want to deviate from revealing the value of $v$ and send a message that
induces an expectation closer to $E^*(v)$. To induce these higher beliefs, the platform needs a large enough jump from the true value. Because the message space is bounded and because there are jumps between beliefs, this means that there is a finite number of induced expectation values possible in equilibrium. The consequence of Lemma 2 is that the true value of $v$ cannot be communicated in equilibrium, i.e., there cannot be full revelation of $v$ in equilibrium.

Given that there is no full revelation, we construct an equilibrium in which the state space is partitioned into $n$ subintervals $[v_0, v_1], [v_1, v_2], ..., [v_{n-1}, v]$, and the platform reveals to the sellers in which interval the realization of $v$ lies. Suppose that the realized state is $v \in [v_k, v_{k+1}]$. Let $m_k$ denote the message sent by communicating a random value drawn from $U[v_k, v_{k+1}]$. Hence, the message $m_k$ can be any value from the interval it represents, which rules out possible out-of-equilibrium beliefs.3

Using Lemma 1, the equilibrium belief that determines the prices will be $E(v|m_k) = \frac{v_{k-1} + v_k}{2}$. To find the boundaries $v_k$ between the subintervals of the message space, we notice that if the true value is $v = v_k$, the platform should be indifferent between sending the messages $m_{k-1}$ and $m_k$. We can write this indifference condition as:

$$\pi^R(v_k, m_{k-1}) = \pi^R(v_k, m_{k}), \; k = 1, \ldots, n-1 \tag{1.13}$$

which can be rewritten as the following difference equation:

$$v_k = \frac{v_{k+1} + v_{k-1} - (1 - 2\alpha)^2}{3 - 2\alpha} \tag{1.14}$$

with boundary conditions $v_0 = v$ and $v_n = \bar{v}$.

3It is sufficient to focus on uniform distributions for the mixing strategies within intervals because for any other set of mixing distributions, the outcomes will be equal.
The unique solution of equation (1.14) is:

\[ v_k = C_1 \lambda_1^k + C_2 \lambda_2^k + v^* \]  \hspace{1cm} (1.15)

where

\[
\begin{align*}
v^* &= 2\alpha - 1 \\
\lambda_{1,2} &= \frac{3 - 2\alpha \pm \sqrt{(3 - 2\alpha)^2 - 4}}{2} \\
C_1 &= \frac{\bar{v} - v^* - \lambda_2^n(v - v^*)}{\lambda_1^n - \lambda_2^n} \\
C_2 &= \frac{\lambda_1^n(v - v^*) - (\bar{v} - v^*)}{\lambda_1^n - \lambda_2^n}
\end{align*}
\]

This unique solution determines the interval boundaries \( v_k \) for messages sent by the platform to reveal information about the value \( v \) and recommend a price.

Once we know how to find the boundaries that determine messages, a second value that determines the equilibrium is the number of intervals \( n \). How large can \( n \) be? As \( n \) becomes larger, we approach full revelation, which was ruled out by Lemma 2. The fact that \( v_{k+1} \) has to be greater than \( v_k \) for every \( k \) allows us to write a condition that determines the maximum \( n \) possible:

\[
\frac{\bar{v} - v^*}{\bar{v} - v^*}(\lambda_1 - \lambda_2) > \lambda_1^n(1 - \lambda_2) + \lambda_2^n(\lambda_1 - 1) \]  \hspace{1cm} (1.16)

These results are summarized in the following proposition.

**Proposition 3.** When there is a natural number \( n^* > 1 \) such that condition (1.16) holds, then there is a price recommendation equilibrium. In this equilibrium \([v, \bar{v}]\) is divided into \( n^* \) subintervals \([v, v_1], [v_1, v_2], \ldots, [v_{n^*-1}, \bar{v}]\), where \( v_k \) is defined by equation (1.15). When \( v \in [v_{k-1}, v_k] \), the platform draws a value from \( U[v_{k-1}, v_k] \) and sends that value as a message to the sellers.

In the price recommendation equilibrium:
• Equilibrium prices with a message from the subinterval \([v_{k-1}, v_k]\) are:

\[
p_R^H|k = \frac{q}{2} + \frac{v_k + v_{k-1} + 2(1 - 2\alpha)}{(5 - 2\alpha)}
\]

\[
p_R^L|k = -\frac{q}{2} + \frac{v_k + v_{k-1} + 2(1 - 2\alpha)}{(5 - 2\alpha)}
\]

(1.17)

(1.18)

• The ex ante expected equilibrium profits are:

\[
\mathbb{E}(\pi_R^P(v)) = \phi \sum_{k=1}^{n^*} \frac{v_k - v_{k-1}}{\bar{v} - \bar{v}} \frac{(3 - 2\alpha)(4(v_k + v_{k-1} + 2(1 - 2\alpha))^2 + q^2(5 - 2\alpha)^2)}{4(5 - 2\alpha)^2}
\]

(1.19)

\[
\mathbb{E}(\pi_R^P) = (1 - \phi) \sum_{k=1}^{n^*} \frac{v_k - v_{k-1}}{\bar{v} - \bar{v}} \frac{(3 - 2\alpha)(2(v_k + v_{k-1} + 2(1 - 2\alpha)) + (5 - 2\alpha)q^2)}{8(5 - 2\alpha)^2}
\]

(1.20)

\[
\mathbb{E}(\pi_R^L) = (1 - \phi) \sum_{k=1}^{n^*} \frac{v_k - v_{k-1}}{\bar{v} - \bar{v}} \frac{(3 - 2\alpha)(2(v_k + v_{k-1} + 2(1 - 2\alpha)) - (5 - 2\alpha)q)}{8(5 - 2\alpha)^2}
\]

(1.21)

• When \(n^* \geq 2\), the platform and the sellers prefer price recommendation to no recommendation (competition). There exists a \(q^*\), such that the platform is better off under price recommendation compared to centralized pricing if and only if \(q > q^*\). There also exists \(\hat{q}_H\) (\(\hat{q}_L\)) such that a high (low) type seller is better off under recommendation than under centralized pricing if and only if \(q > \hat{q}_H\) (\(q > \hat{q}_L\)).

Proposition 3, which is a major result of the paper, shows that whenever there is a natural number larger than 1 for which the inequality in (1.16) holds, it is more profitable for the platform to give recommendations in equilibrium compared to letting sellers compete without a recommendation. Moreover, when the uncertainty \(q\) is high enough, recommendations are more profitable to the platform (and the sellers) compared to centralized pricing. The intuition is that as \(n\) increases, the profit of the platform also increases, which makes rec-
ommendations preferable. When \( q \) is high enough, similarly to the competition case, profits might also increase above the centralized pricing case.

Along with the bubbling (competition) equilibrium, when \( n^* \geq 3 \) there are multiple price recommendation equilibria. These equilibria differ by how coarse the partition of values of \( v \) is. Theorems 3 and 5 of Crawford and Sobel (1982) establish that in a cheap talk game, both the sender and receiver are ex ante better off in an equilibrium with a larger \( n \). Since the conditions of these theorems hold in our model, the profit of the platform and the sellers increases with \( n^* \).

To understand which one of the multiple cheap-talk equilibria might be reasonably played, we apply the no incentive to separate (NITS) criterion of Chen et al. (2008). NITS states that a sender with the lowest type (i.e., a platform that observes \( v = \underline{v} \)) always prefers the cheap-talk equilibrium payoffs than having the receiver (i.e., the sellers) observe the sender’s true type (i.e., the sellers knowing that \( v = \underline{v} \)). Using this criterion, we can prove the following:

**Corollary 2.** The unique equilibrium that satisfies NITS is the equilibrium with the most refined partition, i.e., with \( n^* \) intervals. Consequently, the platform will provide price recommendations rather than choose competition when cheap talk is possible.

To summarize, we have found conditions under which a platform might prefer to let consumers compete with or without price recommendations. These cases are applicable when the uncertainty in the market about seller quality is high enough. An interesting additional finding is that price recommendations are not always beneficial. In many cases they are not credible and will be ignored by the sellers.

After establishing the conditions for which a platform would prefer to provide price recommendations, we deepen the analysis in the following section to understand the impact on sellers, buyers and the market.
1.5. Market Implications

In this section we compare the benefits for sellers and consumers, as well as the equilibrium demand in the different pricing regimes. We start with illustrating the regions of parameters for which the platform or the sellers are better off in the different pricing regimes. Because the inequalities for these conditions have higher order polynomials, we are only able to provide numerical analyses.

Figure 1 shows the regimes in which each player achieves maximum profit, as a function of $q$ and $\overline{\tau}$ when $\alpha=0.45$ and $\underline{\nu} = 5$. We can see a common pattern emerge: when $\overline{\tau}$ is high and $q$ is low, all players prefer centralized pricing (top left); when $\overline{\tau}$ is low and $q$ is high, all players prefer competition (bottom right); when both $q$ and $\overline{\tau}$ are large, recommendation leaves all players better off. The intuition is that $\overline{\tau}$ captures the amount of information the platform has while $q$ captures the amount of information the sellers have. If $\overline{\tau}$ is small, i.e., close to $\underline{\nu}$, there is little variation in the aggregate demand level and demand is consistent. Consequently there is little value to the platform’s information. Because, in addition, low values of $\overline{\tau}$ cannot sustain the recommendation equilibrium, all players are better off if the sellers are allowed to price based on the information they possess. In contrast, if $\overline{\tau}$ is high and $q$ is low, there is little value to the sellers’ private information and the platform can safely ignore it and centralize pricing. Finally, if both sources of uncertainty are relatively strong, the platform should recommend a price, so that the sellers can combine their private information with the platform’s information.

A surprising feature to observe include the differences among the three figures. High type sellers prefer centralized pricing more strongly than low types and even more than the platform. This result is formally stated in the following proposition:

**Proposition 4.** For $\tilde{q}_H$, $\tilde{q}_L$ and $\tilde{q}$ as defined in Proposition 2, and $\hat{q}_H$, $\hat{q}_L$ and $\hat{q}$ as defined in Proposition 3:

- $\tilde{q}_H > \hat{q} > \tilde{q}_L$
Figure 1: Pricing regimes that the platform and the sellers prefer depending on the values of $q$ and $\bar{v}$. Other parameters: $\alpha = 0.45$.,
• \( \hat{q}_H > \hat{q} > \hat{q}_L \)

The proposition shows the counter-intuitive result that high types prefer to relinquish pricing control to the platform, although they have pricing power in competition against low types. The intuition is that when pricing is centralized, prices are equal across types. If one of the sellers is a high type and the other is low, the high type will obtain a large market share and a substantially larger profit than the low type. When the pricing is decentralized, the high type’s advantage is mitigated by the fact that the low type can lower their price to attract more buyers. Centralized pricing can be exploited by the high type to soften competition from price cutters. If one considers the dynamics of pricing on platforms, this implies that the more platforms centralize pricing, the more we might see higher quality players on the platform.

Next, we consider the expected total size of the market, i.e., the expected mass of buyers served in equilibrium. Total demand is equal to

\[
TD(p_1, p_2) = 2(1 - 2\alpha) + 2v + q_1 + q_2 - p_1 - p_2
\]

Using symmetry and integrating over \( q \) and \( v \), the expected total demand is:

\[
E(TD(p_1, p_2)) = 2(1 - 2\alpha) + \bar{v} + \bar{v} - 2E(p)
\]

Using the fact that the expected total market coverage turns out to depend only on expected prices, we can prove the following result:

**Proposition 5.**

- \( E(TD^R) = E(TD^C) > E(TD^{CP}) \).

- The expected distance between a buyer and the seller they purchase from is \( E(\frac{TD}{4}) \).

Proposition 5 shows that total demand is higher when sellers compete on prices. It does not
change if recommendations are feasible or not. This is because the only differences between prices under recommendation and competition is that under recommendation the expectation over $v$ is conditional on the message from the platform. As the sellers have rational expectations and the prices are linear in those expectations, summing over all possible messages and weighting by the message probability yields the same ex ante expectation and hence the same expected price. Under centralized pricing, the prices are higher on average, as the platform internalizes the substitution patterns between the two sellers and therefore faces less elastic overall demand than each seller individually.

One of the important implications of proposition 5 is that when cheap talk (or competition) are more profitable to a platform than centralized pricing, the effect is not coupled with decreased demand, but rather with an increase in market size. As we discuss later, for many young platforms, growth often comes at the expense of profits, but as our results show, these two goals do not necessarily contradict.

Now we consider the consumer surplus (expected utility) of buyers, which we illustrate in Figure 2. Because buyers have an outside option, they are shielded from some of the risk of experiencing a low realization of $v$ or receiving $-q$. In other words, the downside of participating in the market is limited, similar to a financial “call” option. In this case, from an ex ante perspective, buyers prefer a payoff that varies more, as they can capture more of the upside. The more prices reflect the realizations of $v$ and $q$, the less variation there is in the buyers’ payoff. Therefore, buyers prefer those pricing regimes that attenuate the uncertainty the most when translating from realizations of $v$ and quality to prices. Buyers always prefer no recommendation to recommendation, since then prices do not vary with $v$. They also prefer centralized pricing when $\overline{v}$ is small and therefore $v$ matters little, while $q$ is large.

The intuition is formalized in the following result:

**Proposition 6.** When comparing the consumer surplus of buyers under the three pricing regimes:
Figure 2: Pricing regimes that the buyers prefer depending on the values of $q$ and $\bar{v}$ when $\alpha = 0.45$ and

- $CS^C > CS^R$.

- There exists a $q'$ such that $CS^{CP} > CS^C$ if and only if $q > q'$.

The first item of Proposition 6 emphasizes the contradicting preferences of the platform and the sellers with those of buyers. Similarly, buyers prefer centralized pricing only when $q$ is large, which is exactly when the platform and the sellers prefer competition. This result underscores the potential trade-offs between the two sides of the market that platform designers have to consider.

1.6. Impact of Search Technology $\alpha$

An interesting feature of our model is the search technology $\alpha$ that determines what share of consumers see more or less options when visiting the platform. If the platform could influence $\alpha$ by designing a different search algorithm, what would be the platform’s preferred choice?

Performing a complete analysis of how $\alpha$ impacts the equilibria results is non-tractable
because of the complexity of the model. To further the analysis, we therefore use numerical analysis as well as compare the cases of $\alpha = 0$ and $\alpha = 1/2$.

Figure 3 shows the effect of $\alpha$ on the possible number of intervals in cheap-talk equilibria for representative values of $\underline{v}$ and $\overline{v}$. As $\alpha$ increases, more consumers see only one seller, and the platform and the sellers have more aligned incentives. This results in the platform having an incentive to reveal the true value of $v$ more accurately as $\alpha$ increases, up to a point (when $\alpha = 1/2$) in which the platform would reveal the true value of $v$ and the sellers will price without any uncertainty about $v$. A second insight is that when $\overline{v}$ increases, a cheap-talk equilibrium is possible for lower values of $\alpha$, and in such cases, it is more profitable for the platform to select recommendations vs. pure competition.

Finally, we analyze the platform’s preferred choices in the extreme cases of $\alpha = 0$ and $\alpha = 1/2$:

**Proposition 7.** *If the platform can set $\alpha$ to be either 0 or $\frac{1}{2}$ before committing to a pricing regime, it will choose:*

![Figure 3: Number of possible cheap-talk intervals ($n^*$) as a function of $\alpha$ and $\overline{v}$ when.](image)
- *Centralized pricing if* $\alpha = 0$

- *Recommendation if* $\alpha = \frac{1}{2}$

- *The platform always chooses to set* $\alpha = 0$ *and centralize pricing.*

The results show that although the platform would generally prefer price recommendations coupled with limiting choice by consumers, it gains the most when consumers have more choice but the platforms chooses prices for sellers. Because the analysis only focuses on the extreme values of $\alpha$, we are unable to tell whether there is an intermediate value of $\alpha$ in which cheap talk recommendations are preferable to centralized pricing. We leave the question of the interaction of search technology and platform pricing for future research.

1.7. Conclusion

In our analysis, we considered three pricing regimes: (i) competitive pricing by sellers; (ii) centralized pricing by the platform; (iii) recommending prices to sellers. We find that from the platform’s and the seller’s perspective, the optimal choice of the pricing regime depends on the type of uncertainty prevalent in the market. If the aggregate demand uncertainty is more important than the uncertainty about the sellers’ quality, the platform should set prices in a centralized fashion. If the quality uncertainty is larger than the aggregate uncertainty, the platform should let the sellers set their own prices.

A major advantage that the platform can utilize in markets when both types of uncertainties are high are price recommendations. In this case, providing sellers with some information, but not fully revealing it, may increase the profits of the platform above the centralized and the no recommendation case. This increase in profits is not always feasible, as there are cases when price recommendations will not be credible in equilibrium, and sellers will ignore them. Another interesting finding is about which sellers prefer centralized pricing. We found that sellers with high qualities prefer centralized pricing, although intuition would suggest that they would have stronger pricing power and would prefer pricing autonomy.
From the perspective of buyers, competitive decentralized pricing is almost always the best regime. Only when the aggregate uncertainty is small and the quality uncertainty is large, do buyers prefer centralized pricing.

Our analysis uncovers a tradeoff between maximizing the platform profit and consumer surplus which may inform platform designers and managers. Even though we do not model entry of buyers and sellers explicitly, higher expected consumer surplus will often lead to more buyers using the platform and a higher expected seller profit will encourage more sellers to join. Consequently, a growth-stage platform that is willing to sacrifice some profits for larger market share should let sellers set their own prices. A mature platform, in contrast, should use the profit-maximizing pricing regime. In fact, we might interpret the changes in Airbnb’s pricing strategy as following this rule. At first, while the company was growing, Airbnb let the hosts set their own prices. Later they introduced Price Tips, which is a price recommendation service. The introduction of smart pricing takes Airbnb even closer to a centralized pricing system.

The results are of course not without limitations. In order to achieve a tractable solution, we assumed a specific simple demand form. Although we believe the results would hold in more generalized cases, this is still an open question. A second limitation of our model, which would be interesting to explore in future work is the amount of information buyers have, compared to the platform and the sellers. In our model buyers have full knowledge of all relevant model parameters, and relaxing this assumption may be important. Finally, in our game we did not consider entry or exit of the sellers, which is one of the important features that determines platform profits in dynamic platforms such as ride sharing.

In terms of future work, there are two interesting questions that arise naturally from our model and we are considering to focus on. The first is further analysis of the a platform that can design the search technology and pick $\alpha$ to maximize its profit. Platforms often change the amount of search results they display to customers strategically. The second is the impact of the share of revenue the platform takes from sellers on seller behavior. In our
model, because sellers do not enter or exit, this share has no consequence, and extending
the model to capture this effect can be an important next step.

For policymakers, our paper suggests that price recommendation systems may soften com-
petition and potentially harm buyers, compared to not recommending prices. A critical
part of many online platforms’ business model is the status of sellers or service providers as
independent contractors, rather than employees. This allows the platforms to avoid, e.g.,
labor regulation. One criterion for determining the status of an employee vs. a contractor
is their ability to set their own price. Our paper shows that platforms do not always have
to centralize pricing to achieve profits that are above competitive. Price recommendations
allow platforms to extract large profits while avoiding the need to set prices for sellers.
Regulators should therefore consider the impact of price recommendations and its influence
on equilibrium outcomes when they consider the employment status of individuals.
2.1. Introduction

For more and more companies, fairness is becoming a priority in the areas of advertising, targeting, product design and even pricing. Questions of fairness arise whenever limited resources are allocated between different entities. There are three ways in which an allocation procedure can be considered fair: (i) procedural, i.e., whether the same rules and standards are applied equally to all entities; (ii) distributive, i.e., whether the resulting distribution of resources is equal or whether any inequality of outcomes is justified; (iii) retributive (or restorative), i.e., whether the allocation corrects past injustices, either by rewarding the victims or by punishing the perpetrators. While much of the extant literature focuses on procedural fairness (discrimination being a prominent example), in this paper I concentrate on distributive and restorative fairness. Looking through this lens, it is easy to see that the issue of fairness in marketing goes beyond machine learning and algorithms, which has gotten the most “popular press” today as being potentially unfair (i.e., algorithmic bias, Lambrecht and Tucker (2019); Lepri et al. (2018)). Since almost any marketing decision is bound to redistribute resources between the company, its competitors, consumers, employees, suppliers etc, no matter how it is made, the question of fairness is always relevant. But fairness can be at odds with the profit-maximization imperative of the firm. Many managers regularly face a dilemma between efficiency (profit-maximization) and fairness.

I study the trade-off between fairness and efficiency using data from the legal cannabis market in the state of Washington. One of the major concerns surrounding cannabis legalization, expressed by academics and activists, is that since African-American communities bore the brunt of the social costs of the war on drugs in the form of mass incarceration and increased policing, African-Americans should be able to benefit the most from legal sales of cannabis. These groups argue that benefits should come from both business ownership and
consumer access. John Hudak writes for the Brookings Institution:

The future of cannabis policy in the United States, however, must include expungement (preferably, automatic expungement), but also more comprehensive efforts to help the communities that have been ravaged by the War on Drugs.

...more effective policies must be implemented in legalizing states to create new and lasting ownership opportunities for people of color and those with previous, low-level cannabis convictions.¹

These are considerations of distributive (who benefits from legalization) and restorative (will the legalization correct past wrongs) fairness.

However, the text of the ballot initiative that legalized cannabis (I-502) does not mention racial justice as a primary goal of the reform, instead stating explicitly that the aim of the new approach to cannabis is to “generate new state and local tax revenue for education, health care, research, and substance abuse prevention”.² Washington state policymakers found themselves making a choice similar to the one that many marketing managers currently have to consider. On the one hand, there was an explicit mandate to implement the reform in a way that raises the most money for the state. On the other hand, the public demanded a fair distribution of licenses.

What allocation mechanisms could the WA state government choose to ensure either efficiency or fairness? There is a large literature in economics that argues that auctions are typically the best way to allocate resources in the most efficient manner. Auctions are used by many governments to award contracts and licenses, including the Washington state government that uses auctions to allocate logging rights, so there was certainly no lack of expertise to implement this mechanism. However, the state government chose to

¹“Marijuanas racist history shows the need for comprehensive drug reform”, https://www.brookings.edu/blog/how-we-rise/2020/06/23/marijuanas-racist-history-shows-the-need-for-comprehensive-drug-reform/, accessed on 1/4/2021

²The other two express goals are focused on law enforcement: “allows law enforcement resources to be focused on violent and property crimes” and “take marijuana out of the hands of illegal drug organizations and brings it under a tightly regulated, state-licensed system similar to that for controlling hard alcohol”
use a lottery to allocate licenses. Under the lottery mechanism selected by the state of Washington, every applicant has an equal chance of winning, so this mechanism is fair in both a procedural and distributive sense. One might argue that if the state was concerned about providing opportunities for minority entrepreneurs, they should have targeted them explicitly, either by setting a quota or using a scoring mechanism (e.g., an auction, in which a bid is adjusted by a factor that is a function of demographics of the bidder). While this would potentially be the best way to achieve restorative fairness, there are two immediately identifiable practical problems with this method of allocation. First of all, note that such a mechanism is not procedurally fair. Second, this mechanism is likely to be politically contentious, which would at best dramatically prolong the allocation process and at worst imperil the legalization itself. Illinois allocation program is a helpful example. When Illinois voted to legalize cannabis in 2019, seven years after Washington state had done so, the legislation explicitly posited restorative justice as one of its goals. Illinois, too, chose to allocate cannabis retail licenses by lottery. Even with restorative justice as a driving motive for the legislation, however, the only revision they made to Washingtons allocation model was to make it easier for minority-owned businesses to qualify for the lottery. A lottery is the fairest mechanism for cannabis license allocation given the current political constraints.

Washington state's cannabis market presents a unique opportunity to quantify the trade-off between fairness and efficiency due to the availability of very detailed data on both the applications and market transactions after the lottery. The first goal of this paper is to find the “price of fairness”, i.e. how much money did the WA state government leave on the table by choosing a more equitable allocation mechanism. The “price of fairness” consists of two parts: the auction revenue and the foregone sales tax revenue the government would have received by virtue of selecting more efficient retailers into the market. The second goal is to document whether the efficiency gains from the auction are distributed fairly across demographic groups. To achieve these goals, I construct a model of competition in the recreational cannabis market.
Since I do not observe the counterfactual auction outcomes, I take a structural approach and explicitly model the auction allocation mechanism. In my model, firms bid for licenses under complete information about their competitors. The winners of the auction then enter the market and compete in prices. The value of a license for each retailer then depends on their own characteristics and the characteristics of their competitors. Because of the complete information assumption, this game has multiple equilibria. I overcome this issue by a selection mechanism based on the assumption that the firm with the highest average profitability will be the most likely to win.

I estimate an equilibrium model with nested logit demand and Bertrand-Nash pricing using detailed transaction data from the Washington State Liquor and Cannabis Board (WLCB) for the period from August 2014, when the first retailers entered the market, to May 2017. I then use the estimated model to compute retailers’ profits under counterfactual unobserved license allocations. Given these profit functions, I find the counterfactual allocation under the auction. Next, I compare prices, quantities and tax revenues between the auction outcome and the lottery. The difference in tax revenue, combined with the auction revenue, constitutes the price of fairness.

I allow firms to be heterogeneous along three dimensions: (i) cost efficiency, (ii) ability to sell (quality of service) and (iii) location. This means that two different firms (i) face different marginal costs (ii) face different demand due to differences in quality (iii) sell to a different population of consumers and face different spatial competition. Location affects demand as consumers prefer to shop closer to where they live. I get the addresses for all applicants from the WLCB application data, which I use as the potential store locations. However, I do not observe cost and demand efficiency for retailers who did not win the lottery. To overcome this issue, I estimate the joint distribution of demand and cost types for the observed retailers and assume that the types of unobserved retailers come from the same distribution, conditional on the county. Note that due to the lottery, the set of observed retailers is a representative sample from the set of applicants. In other words, given a fixed
set of applicants for licenses, we do not have to adjust for selection (a la Heckman (1979)) when estimating the distribution of retailer types.

A common issue when estimating demand is price endogeneity (Villas-Boas and Winer (1999)). I use three sets of instrumental variables to overcome this problem: (i) lagged rainfall and temperature, which affect the production costs of cannabis; (ii) wholesale prices in other markets; (iii) average upstream prices. I find a median price elasticity of 1.57. I also estimate the elasticity with respect to distance of 0.19, so a 1% increase in distance is equivalent to a 0.12% increase in price. Finally, I estimate the aggregate elasticity for the cannabis category as a whole of 0.96, which implies that most substitution happens within the cannabis category.

For the counterfactual results to be credible, it is crucial that my demand model and my approach to unobserved retailers provide me with accurate predictions of demand for those retailers and their competitors. Fortunately, the data provide an opportunity to validate my model. Since not all firms enter the post-lottery market at the same time, I often observe changes in market structure. In 157 instances, I observe a market before and after an entry occurs. I select those cases and predict the post-entry equilibrium as if I do not observe the entrant’s type. Comparing the predicted and realized market outcomes, I can get a sense of my models accuracy. In terms of quantity, my model underpredicts by 17% on an individual store level and only by 3% on market level. For prices, the error is smaller: 3% on individual firm and 1% on market level. Since I am mostly concerned with market outcomes, I conclude that the model performs reasonably well.

Using the auction counterfactual, I find that compared to a lottery, an auction selects on average more efficient retailers (with lower marginal costs and higher quality). The auction leads to a modest increase in total quantity sold (5%). The reason that it is not even higher (more efficient) is the high degree of substitution across products and retailers within the same market, so most of the increase in demand for an individual retailer comes from business-stealing, rather than market expansion. As a result, price competition intensifies.
(prices are 3% lower under auction on average). Since the increase in quantity is higher than the decrease in prices, the sales tax revenue is 2% higher under an auction design. According to my counterfactual, the government loses more than $8M annually in sales tax revenue because of the choice to use a lottery. Additionally, the auction revenue is estimated to be between $57.3 and $63M. Assuming a discount rate of 1%, the price of fairness over a ten year horizon is more than $137M or approximately 0.39% of the state’s annual budget.

I use a machine learning algorithm by Sood and Laohaprapanon (2018) to predict applicants’ race from their names, which I collect from the Washington State Department of Labor and Industries. Combining those predictions with my counterfactual results, I find that compared to a lottery, an auction would lead to a loss of 10 licenses for Black applicants and 1 license for Hispanic applicants, which is a reduction of 21% and 10% respectively.

My model allows me to estimate consumer surplus on census block group level under an auction and a lottery. I combine these estimates with the census demographic data on racial composition of the block groups. I find that (i) majority-White block groups’ consumer surplus increases by 20%, while majority-Black block groups’ increases only by 3% and (ii) consumer gains from an auction are increasing in the population share of White residents, from 11% gain for the least White block groups to 32% for the most White. This racial disparity cannot be explained by income differences, as the consumer gains decrease in median household income.

I also investigate how much unfairness increases due to auction. First, I document a substantial widening of the gap between the retailers who are most likely and least likely to receive a license when an auction is used. Second, I try to see if this inequality spills over to the consumer side: are welfare gains geographically concentrated? I find that (a) rural areas on average gain more from an auction than urban areas; (b) overall geographic inequality (on census block group level) increases by 3%.

I find that the trade-off between fairness and efficiency is real and binding: a switch from
a lottery to an auction produces substantial gains in state revenue and consumer surplus, but also increases the inequality in ex ante probability of winning between firms and spatial inequality between consumers. These findings are relevant to policymakers dealing with questions of fair regulation, and to platform companies considering incorporating fairness into their market design.

2.1.1. Related Literature

There are multiple streams of literature that this paper contributes to. First, there is a growing literature in marketing, economics and computer science recognizing that the issues of algorithmic fairness go beyond technical issues such as e.g. biased training data, and cannot be solved by only developing unbiased algorithms. Rambachan et al. (2020) argue for a larger role of economic analysis in the study of algorithmic fairness. Ali et al. (2019) and Lambrecht and Tucker (2019) argue that cost optimization may skew the display of broadly targeted ad campaigns if different demographics are priced differently. Nasr and Tschantz (2020) develop bidding strategies for online auctions for advertisers to avoid results biased by gender and to quantify the efficiency loss for bidders from those strategies. I extend this literature by considering a fully egalitarian allocation mechanism on the part of the platform (government) and quantifying its subsequent loss to the platform.

Second, there is a large body of theoretical work on auctions with aftermarkets, i.e. auctions with downstream interactions between bidders (Jehiel and Moldovanu (2000)). Goeree (2003), Varma (2002) consider bidding in such an auction as a signaling device and characterize how different auction mechanisms affect signaling behavior. Janssen and Karamychev (2009, 2010) study the conditions under which auctions with aftermarkets do not allocate the licenses to the most efficient bidders. Their key insight is that even if ex ante bidder types are independent, conditional on winning, there is a positive correlation between bidder types. If an efficient bidder wins, they expect other winners to be even more efficient and vice versa. If competition between inefficient bidders is soft enough, they may bid higher than efficient bidders. In this case, an efficient equilibrium does not exist. In my paper, I
simulate a counterfactual auction using the profit functions estimated from the data. To the extent of my knowledge, this is first empirical paper focusing on auctions with aftermarkets.

Third, there is a stream of empirical literature studying allocation problems in general and auctions in particular (Hendricks and Porter (1988)). Particularly relevant are the studies which model post-allocation behavior by winning bidders. Bajari et al. (2014), An and Tang (2019) study the allocation of incomplete procurement contracts and the hold-up problem between the government and the contractors. Another relevant stream of literature is the empirical work on multidimensional screening and scoring auctions (Krasnokutskaya et al. (2020), Lewis and Bajari (2011)). In these papers, similar to mine, the seller (the principal) has preferences over not just the cost of the project, but potentially other variables (time to completion, quality, cost overruns). If these variables could be written into a contract, a scoring auction can be used. Otherwise, the contract is incomplete and there is a hold up problem. This paper adds to this literature by considering that the principal may have preferences not only over the outcomes (e.g. revenue), but also over the characteristics of the allocation mechanism itself (fairness).

Fourth, as cannabis legalization is becoming more common in the United States, many scholars turned their attention to the legal cannabis market. Hollenbeck and Uetake (2019) study the pass through rate of the sales tax in the legal cannabis market in Washington and construct a Laffer curve for the tax. In Escudero (2019), the author considers the loss cannabis retailers experience from using rule-of-thumb pricing instead of profit maximization.

The paper closest to mine is Thomas (2019). In her paper, she considers the effect of the license cap on the legal cannabis market in Washington state. Using a supply and demand model, she simulates the market under a free entry regime and finds that it would boost the overall surplus by 18%. To separate the effect of license cap from the effect of location randomization, she then runs an auction counterfactual. The key difference between her paper and mine is the focus on an entry model versus the auction. In my paper, I focus on
capturing all possible heterogeneity among retailers that could influence their auction bids. I investigate the efficiency of auction mechanism and the level of inequality it produces.

2.2. Background and Data

In November 2012, the state of Washington passed a ballot measure legalizing regulated production, sale and consumption of recreational cannabis. The law created a three-tier licensing system for producers (i.e. growers), processors (i.e. wholesalers who buy raw cannabis and either repackage it into smaller retail amounts or convert it into pre-rolled joints, edibles, vapes etc) and retailers.

The state was divided into 123 jurisdictions (city or town or rural area of a county outside major cities). Retail licenses are capped on the level of jurisdiction. For example, in Spokane county, out of 31 total licenses, 14 are assigned to Spokane, 5 to Spokane Valley and 12 to the rest of the county. The total number of licenses was originally capped at 334 and then in January 2016 increased to 556. To apply for a license, the applicant had to submit the proposed address for the retail operation (to check for compliance with state law regarding proximity to schools etc) and pay an application fee of $250. In 75 out of 123 jurisdictions, the number of applicants exceeded the number of licenses and the state had to run a lottery. These are the markets that this paper focuses on, as in the markets without oversubscription the licenses are not contested. Retailers are not allowed to be vertically integrated and there are restrictions on horizontal integration (one entity cannot hold more than 3 retail licenses or more than a third of all licenses in a given jurisdiction).

Because cannabis was and still remains an illegal drug on the federal level, the policymakers in Washington were very concerned about potential diversion of cannabis across state lines. To prevent that from happening, the state implemented a bio-tracking system that requires the producers, processors and retailers to log every operation with cannabis from the moment a seed is planted to the final retail sale. Each plant has a unique ID, which is then attached to the cannabis products (e.g. edibles) that are made from it. Therefore,
the bio-tracking data contains records of all transactions on all levels of the legal cannabis market in Washington: weight, transaction price, product type, identities of the transacting parties, date and time of transaction.

Cannabis is not a standardized product, so there are no SKUs or equivalents. The dataset provides 13 product categories, which I further aggregate to 5: usable cannabis (parts of the plant, usually sold by weight or as a pre-rolled joint), solid edible, liquid edible, vape, and other, which combines multiple small categories, such as tinctures or patches. I treat these categories as homogenous products, i.e. I ignore the differences between different strains or different types of solid edible (chocolate vs cookies).

Figures 4 and 5 show the historic sales and average prices by category. The vast majority of sales are for the cheapest product, which is usable cannabis. There is a general increasing trend in sales, which slows down somewhat towards the end of the sample. It is accompanied by a downward trend in prices. This picture is consistent with more retailers entering throughout the sample period, even though their total number is restricted by the license cap. Perhaps more importantly, there is no cap on the number of licenses for cannabis producers (growers). There is an increasing number of producers entering the market throughout the sample period, sending the wholesale prices further and further down. Figure 6 presents the time series of retailer margins. Margins are computed using the wholesale prices, i.e. $\zeta = \frac{p - w}{w}$, where $\zeta$ is the margin, $p$ is the retail price and $w$ is the wholesale price. After the initial volatile period, retail margins are fairly stable over time, with only a slight downward trend. This implies that the downward price trend is mostly driven by the downward trend in wholesale prices and not by increased competition or learning.

For the purposes of the auction counterfactual, I need to establish that there is in fact heterogeneity among retailers. In Figure 7, I present the distribution of average margins by retailer. They are both unusually high for a retail sector (65% on average, compared to, for example, 30-50% in grocery retail) and very dispersed. This suggests that retailers are
Figure 4: Historical growth of cannabis retail sales in Washington

Figure 5: Time series of cannabis retail prices in Washington
heterogeneous in how they operate.

The state did not collect data on race of applicants. However, each applicant received a Unique Business Identifier (UBI), which can be used to look up business information on the website of Washington State Department of Labor and Industries, including the name of the business owner. This way, I collected names of 1170 applicants, which accounts for 73.9% of all applicants. The missing names are companies that were dissolved for a long enough time for records to be deleted. The dataset contains the names for all retailers that won the lottery. I use a machine learning algorithm by Sood and Laohaprapanon (2018) that predicts a person’s race based on first and last name. They use Florida voter registration data as the training dataset. For each name, their model gives me probabilities of that person belonging to one of 4 demographic groups: non-hispanic white, Hispanic, non-hispanic Black and Asian.

Table 1 presents two sets of summary statistics. First, for the lottery winners, I present the market outcomes: average monthly sales in grams, average prices (quantity-weighted) and assortment size (average number of categories offered). Second, for the entire set of
applicants, I present the characteristics of the area (2010 census tract) where the business is located: density in ppl per square mile, median annual household income in dollars, and median age of residents in years.

Since I only have probabilistic estimates of business owners’ racial identities, I compute the relevant variables for all retailers and applicants and then compute an average weighted by the probability a given retailer belongs to group $g$:

$$\bar{Y}_g = \frac{\sum_{r \in R} Y_r \mathbb{P}_{rg}}{\sum_{r \in R} \mathbb{P}_{rg}}$$  \hspace{1cm} (2.1)

where $Y_g$ is the variable of interest (such as sales) conditional on owner being in group $g$, $R$ is the set of all retailers or applicants, $\mathbb{P}_{rg}$ is the probability retailer $r$’s owner belongs to group $g$.

Black retailers have the highest average sales volume and the lowest prices. Non-hispanic white retailers have substantially higher prices, but their sales are not much lower. Asian retailers have the lowest sales despite having second-lowest prices. Assortment sizes do not vary much between groups, with almost everybody carrying all product types. Most
Table 1: Retailer characteristics by race of owner. For lottery winners: monthly cannabis sales in grams, average (quantity-weighted) price, assortment size (number of product types available in an average month). For all applicants: density of the Census Tract in ppl/mi², median household income in the census tract, median age in the census tract.

<table>
<thead>
<tr>
<th></th>
<th>Monthly sales</th>
<th>Price</th>
<th>Assortment Size</th>
<th>Area Density</th>
<th>Area Income</th>
<th>Area Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>asian</td>
<td>13734.1</td>
<td>12.2</td>
<td>4.8</td>
<td>3981.08</td>
<td>48726.89</td>
<td>36.68</td>
</tr>
<tr>
<td>hispanic</td>
<td>14852.0</td>
<td>12.9</td>
<td>4.7</td>
<td>3228.70</td>
<td>49785.77</td>
<td>36.91</td>
</tr>
<tr>
<td>black</td>
<td>17357.0</td>
<td>11.9</td>
<td>4.8</td>
<td>3618.71</td>
<td>50575.46</td>
<td>38.21</td>
</tr>
<tr>
<td>non-hisp. white</td>
<td>16268.4</td>
<td>12.7</td>
<td>4.8</td>
<td>3302.85</td>
<td>53028.61</td>
<td>38.23</td>
</tr>
<tr>
<td>no name</td>
<td>3370.39</td>
<td></td>
<td></td>
<td>3387.61</td>
<td>53329.52</td>
<td>38.13</td>
</tr>
<tr>
<td>Average</td>
<td>16031.4</td>
<td>12.6</td>
<td>4.8</td>
<td>3387.61</td>
<td>53329.52</td>
<td>38.13</td>
</tr>
</tbody>
</table>

applicants tend to be in areas that are relatively dense, like cities and towns (for Washington state as a whole, population density is 100 people per square mile, for King county 1000 people per square mile). Black and Asian applicants tend to be in denser and poorer areas.

2.3. Model

2.3.1. Simplified Model

I start by analytically solving a very simplified version of my model to illustrate some of the main forces. Suppose there is only one license to be allocated, so the winner of the lottery or auction is going to be a monopolist. There is a population of firms applying for the monopoly license. Their heterogeneity is captured by a two-dimensional type \((\eta, c)\), where \(\eta\) combines quality of the product, management and location and \(c\) is the marginal cost. Suppose the demand in the market is linear: \(D(p) = \eta - p\). Then the monopolist maximizes profit \(\pi^m = (p(1 - \tau) - c)(\eta - p)\), where \(\tau\) is the sales tax and the monopoly price is,

\[
p^m = \frac{\eta(1 - \tau) + c}{2(1 - \tau)}
\]

with corresponding monopoly profit equal to \(\pi^m = \frac{(\eta(1 - \tau) - c)^2}{4(1 - \tau)}\).

Under an auction, the firm with the highest profit in the product market wins. In other words, for any draw of applicant types, the winner is the firm with the highest \(\eta(1 - \tau) - c\). Total quantity sold in this market is \(Q = \frac{\eta(1 - \tau) - c}{2(1 - \tau)}\), which implies that the auction selects
the firm that provides the highest quantity. This means that the consumer surplus is also
the highest under an auction design.

However, as per equation 2.2 above, price is a function of the sum \( \eta(1 - \tau) + c \), rather than
the difference. Then the prices can be higher or lower under the auction. Prices increase
under an auction, for example, if the costs are a fraction of quality, i.e., when the two types
are positively correlated. The tax revenue is \( \tau \frac{\eta(1-\tau)^2-e^2}{4(1-\tau)} \). Once again, the effect of the
auction is ambiguous. If the price goes down while the total quantity increases, the tax
revenue may still go down.

2.3.2. Product market

The government is allocating \( L \) licenses to \( N \) potential retailers. If \( L \geq N \), then each retailer
gets a license. If \( N > L \), there is either an auction or a lottery to determine who gets the
licenses. There are \( \min\{L, N\} \) retailers in the market. In every period \( t \), retailer \( r \) sets the
vector of prices \( p_{rt} \) for \( J \) products. Demand for product \( j \) at retailer \( r \) at time period \( t \),
\( D_{jrt}(p_t) \), is a function of the price vector. Then the realized variable profit in period \( t \) is

\[
\pi_{rt} = \sum_{j=1}^{J} (p_{jrt}(1 - \tau) - c_{jrt})D_{jrt}(p_t) \tag{2.3}
\]

where \( c_{jrt} \) is the marginal cost of product \( j \) in period \( t \) for retailer \( r \) and \( \tau \) is the sales tax.
The retailer is maximizing profit by choosing prices for all products it is offering, taking
into account the substitution patterns between their own products, the products of their
competitors, and the outside good.

I assume that retailers set prices in a Bertrand-Nash fashion, i.e. retailer \( r \)'s prices are
optimal given the prices of all other retailers present in the market. Then the following
system of FOCs holds for the vector of equilibrium prices $p^*_t$:

$$D_{jrt}(p^*_t)(1 - \tau) + \sum_{k=1}^{J}(p^*_{krt}(1 - \tau) - c_{krt})\frac{\partial D_{krt}(p^*_t)}{\partial p_{jrt}} = 0 \quad j = 1, ..., J \quad r = 1, ..., \min\{N, L\} \tag{2.4}$$

Denote by $\pi^*_{rt}$ the equilibrium profit.

### 2.3.3. Demand specification

I parameterize the demand as a nested logit. The utility a consumer $i$ gets from purchasing product $j$ from retailer $r$ at time $t$ is:

$$u_{ijrt} = \beta_j - \alpha p_{jrt} + \eta_r + \mu_t + (\phi_1 + \phi_2 * \text{density}_i)d_{ir} + \xi_{jrt} + \rho \bar{\varepsilon}_{ih(j)rt} + (1 - \rho)\varepsilon_{ijrt} = \nu_{jrt} + (\phi_1 + \phi_2 * \text{density}_i)d_{ir} + \rho \bar{\varepsilon}_{ih(j)rt} + (1 - \rho)\varepsilon_{ijrt} \tag{2.5}$$

where

- $\beta_j$ is a product type fixed effect, which captures preferences consumers have over different products
- $\alpha$ is price sensitivity, i.e. disutility from paying money for the products
- $\eta_r$ is retailer fixed effect, which captures the retailer-specific characteristics, such as customer service or assortment within a specific product type
- $\mu_t$ is a time fixed effect, which captures fluctuations in demand by time period common across markets (e.g. holidays when people have more time off work, so they can consume more cannabis)
- $d_{ir}$ is distance between the retailer and the consumer and $\phi$ is the disutility of traveling a mile
- the disutility of traveling is higher for urban areas than for rural areas, so it is
creasing in population density around consumer $i$

- $\xi_{ijrt}$ is an unobserved transitory shock, specific to a particular retailer, product and time. This shock combines all unobserved factors that may affect demand at a particular retailer at a given time, which we do not observe, i.e. weather at the retailer’s location, local events etc. I assume that while these shocks are not observed by me, they are observed by retailers and taken into account when setting prices.

- $\varepsilon_{ijrt}$ is individual product-level shock and $\bar{\varepsilon}_{ih(j)rt}$ is the nest-level shock, i.e. any unobserved factors that increase or decrease consumer $i$’s utility of consumption of cannabis of a particular product or a group of products (nest).

- $V_{jrt}$ is the deterministic part of the utility that is also common for all consumers (i.e. “average” utility of product $j$ at retailer $r$ in month $t$).

- $\rho$ is a parameter capturing the level of substitution within the nest. If $\rho = 0$, the model reverts to standard multinomial logit, in which Independence of Irrelevant Alternatives (IIA) holds for all products, including the outside option. If $\rho$ is close to 1, all substitution between products happens within the nest. IIA holds within the nest, but not between nests.

Assuming that $\varepsilon_{ijrt}$ and $\bar{\varepsilon}_{ih(j)rt}$ are type-I extreme value and normalizing the deterministic component of the outside option to zero, I get the following purchase probability:

$$s_{ijrt} = \frac{\exp((V_{jrt} + (\phi_1 + \phi_2 \ast density_i)d_{ir})/(1 - \rho)) \exp(V_{ih(j)rt}/(1 - \rho))}{1 + \sum_h \exp(V_{ihrt})}$$  \hspace{1cm} (2.6)

where $V_{ih(j)rt}$ is the inclusive value of the nest $h(j)$, which is the nest to which product $j$ belongs:

$$V_{ih(j)rt} = (1 - \rho) \log \sum_{k \in h(j)} \exp \left[ \frac{V_{krt} + (\phi_1 + \phi_2 \ast density_i)d_{ir}}{1 - \rho} \right]$$  \hspace{1cm} (2.7)
and the demand for product $j$ at retailer $r$ in period $t$ is

$$D_{jrt} = M s_{jrt} = M \int s_{ijrt} d\mathbb{P}(i)$$

(2.8)

where $M$ is the market size and $\mathbb{P}(i)$ is the measure that captures the spatial distribution of consumers in the market.

To proceed, I further parametrize the demand model. A time period is a month. I take each county as a separate market. I use the 2010 census information on population by block group to estimate the spatial distribution of consumers. Since we only observe consumers on a block group level, the distance I use is the distance from the centroid of the block group to the retailer $r$’s location. So the market demand then is estimated as

$$\hat{D}_{jrt} = M \sum_{b \in C} w_b s_{bjrt}$$

(2.9)

where $w_b$ is the population weight of block group $b$, $C$ is the set of all block groups in the county. Even though the market is a county as a whole, I assume that competition for each retailer is localized. To do this, for each block group, I find the five closest retailers (within the county) and assume that consumers do not visit any other retailers. This way a retailer’s demand is not affected by a price change in a retailer that is several hours drive away.

I assume that the market size is 6 times the adult population of the county.\footnote{This is based on the back-of-the-envelope calculation that an average person needs 0.2 grams of usable cannabis to get intoxicated, so if every adult gets intoxicated every day of the month, the total consumption is 6 grams per adult every month.} I use the nesting structure to capture the substitution between the inside and the outside good, so all products within a market are in the same nest.
2.3.4. Cost specification

I parametrize the marginal costs in the following way:

\[
\log(c_{jrt}) = \lambda_j + \omega_r + \psi_t + \nu_{jrt} \tag{2.10}
\]

where

- \(\lambda_j\) is the product fixed effect, as different product types have different wholesale prices and require different costs, for example, edibles are stored differently from vapes
- \(\omega_r\) is a retailer fixed effect, which captures differences between retailers in their ability to negotiate with wholesalers or in general run the business more efficiently
- \(\psi_t\) is a time fixed effect, as the wholesale prices may differ depending on time of the year (e.g. lower during the harvest season).

2.3.5. Counterfactual allocation mechanism

In this subsection, I describe the proposed counterfactual allocation auction mechanism.

I assume that the retailer cost and demand types are common knowledge among retailers, even though they are not observed by the econometrician. This means that the auction stage is a full information game.

Denote by \(\Omega_m\) the set of retailers in market \(m\). The equilibrium profit of retailer \(r\) is a function of \(\Omega_m\): \(\pi_{rt} = \pi_{rt}(\Omega_m)\). Denote by \(\pi_r(\Omega_m)\) the total discounted future profit of retailer \(r\) net of fixed costs. I assume that retailers have perfect foresight, they know the types of their competitors and they do not anticipate any additional entry in the future. Then when they are bidding for the licenses, they have no uncertainty regarding \(\pi_r(\Omega_m)\).

I assume that the fixed costs are identical for all retailers within a market. This is a strong assumption. On the one hand, it is reasonable to assume that many of the typical
costs of running a small business (rent, prevailing wages, utilities etc) would be similar between retailers in the same market. On the other hand, it is also easy to imagine that the retail stores are run very differently, which leads to differentiation in fixed costs, which is potentially correlated with the retailer’s quality. Weakening this assumption is subject of future work.

The allocation mechanism I consider is a uniform L+1st price auction. Since all licenses are the same, any non-uniform price mechanism would be not envy-free and therefore not stable, since some firms would pay more for the same license. L+1st price (i.e. everybody pays the highest losing bid) guarantees that every market participant receives a positive surplus in equilibrium.

Denote by $b^*_m$ the equilibrium payment in market $m$. Then the Nash equilibrium in market $m$ is a pair $(\Omega^*_m, b^*_m)$. Denote the identity of the winner with the lowest profit as $r^* = \arg\min_{r \in \Omega_m} \{\bar{\pi}_r(\Omega^*_m)\}$. Then the allocation is an equilibrium if

$$\bar{\pi}_{r^*}(\Omega^*_m) \geq b^*_m \geq \max_{r' \notin \Omega^*_m} \{\bar{\pi}_{r'}(\Omega^*_m \setminus r^* \cup r')\} \quad (2.11)$$

In other words, as long as none of the retailers that are not in the market would choose to trade places with the least profitable retailer that is in the market, the allocation is an equilibrium. First, note that the equilibrium payment $b^*_m$ and therefore the state’s auction revenue is not pinned down, but it is bounded from above and below. Going forward, I will report the maximum and the minimum of auction revenue.

Second, there are potentially multiple equilibrium allocations, i.e. $\Omega^*_m$ is not guaranteed to be unique. When I compute the auction counterfactual in section 2.6.1, I use a procedure that mimics a descending auction or a sequence of single-license auctions. The key idea is that the first person to enter would be the retailer which is best off independent of the specific allocation. In entry literature, it is often assumed that firms enter the market in order of profitability (e.g. Berry (1992)). My procedure serves the same purpose (pinning
down a unique equilibrium) and follows a similar logic (the highest profitability firms are most likely to be “leaders”).

More formally, I go throw the following steps:

1. Compute $E(\pi_r(\Omega)) \ \forall r$ assuming all possible realizations of $\Omega$ are equally likely

2. Pick $r_1 = \arg\max_r E(\pi_r(\Omega))$, i.e. the retailer with the highest a priori expected profit

3. Compute $E(\pi_r(\Omega)|r_1 \in \Omega)$ assuming all possible realizations of $\Omega|r_1 \in \Omega$ are equally likely

4. Pick $r_2 = \arg\max_r E(\pi_r(\Omega)|r_1 \in \Omega)$

5. Compute $E(\pi_r(\Omega)|\{r_1, r_2\} \subset \Omega)$ assuming all possible realizations of $\Omega|\{r_1, r_2\} \subset \Omega$ are equally likely

6. Repeat steps 4-5 until all licenses are allocated.

I then verify that the resulting allocation is in fact a Nash equilibrium. Note that at this stage I do not take expectation over $\Omega$, since the distribution is degenerate (containing only the candidate allocation with probability 1).

2.4. Estimation

In this section I discuss my approach to estimating demand, marginal and fixed costs. Then I talk about identification and summarize the estimation results. I conclude by conducting a validation exercise to understand whether the model can do a good job predicting demand for retailers that I do not observe.

I estimate the model by GMM using the standard Berry et al. (1995) algorithm with a few computational improvements taken from Conlon and Gortmaker (2019). Note that I can compute the market shares of all products for a given vector of homogeneous utilities $V = \{V_{jr}\}$ and non-linear parameters $\theta_2 = (\rho, \phi_1, \phi_2)$ using equation (2.6). Call these
market shares \( \hat{s}_{jrt}(V, \theta_2) \). I can find a vector of homogeneous utilities \( V(\theta_2) \) such that the computed market shares are equal to the one observed in the data: \( s_{jrt} = \hat{s}_{jrt}(V, \theta_2) \). This is a non-linear system of equations that does not have a closed-form solution. Thankfully, it can be solved using a simple iterative algorithm which is shown to be a contraction mapping (Berry et al. (1995), Grigolon and Verboven (2014)):

\[
\log V^{i+1} = \log V^i + (1 - \rho)(\log s - \log \hat{s}(V^i, \theta_2)) \tag{2.12}
\]

where the \( i \) superscript denotes the algorithm’s iteration. The algorithm stops when \(|V^{i+1} - V^i| < 10^{-8}\).

Note that if \( \rho \) is close to 1, equation (2.12) leads to very slow updating. To avoid this problem, I use an accelerated version of the fixed point algorithm, SQUAREM, which is essentially a numerical approximation of Newton-Raphson method.

Given the vector \( V \), I can estimate all linear parameters using 2SLS to account for price endogeneity:

\[
V_{jrt} = \beta_j - \alpha_p p_{jrt} + \eta_r + \mu_t + \xi_{jrt} \tag{2.13}
\]

The moment condition is the following:

\[
\mathbb{E}(\xi Z) = 0 \tag{2.14}
\]

where \( Z \) is the matrix of instruments, discussed in the next section.

I use the residuals \( \hat{\xi}_{jrt}(\theta_2) \) from (2.13) to construct the GMM objective function:

\[
g = \hat{\xi}' Z W^{-1} Z' \hat{\xi} \tag{2.15}
\]

where \( W \) is the weighting matrix. First I run the optimization using an identity matrix and then I use the moments from the first step to estimate the efficient weighting matrix and use it in the second step of the estimation.
2.4.1. Identification

As one can see from equation (2.4), I assume that retailers take $\xi_{jrt}$ into account when making pricing decisions. Therefore, the prices are endogenous. I use three sets of instruments, all three of which try to capture the upstream cost shocks that should shift retailers’ marginal costs, but not the demand.

1. Weather: lagged (by month) temperature and rainfall. Cannabis is an agricultural product and a substantial portion of it is grown outdoors, so weather affects the size of the harvest, which in turn affects the prices. I use the county-level weather information, which is matched to the location of cannabis producers (growers). Since every plant has a unique id number, I can trace them from producer to the final sale. Then I take a weighted average of the weather variables for each product sold. In other words, I take a weighted average over the producers whose products the retailer sold in a given month. I take lags both to avoid potential effects of weather on demand and as a reflection that past, not current, weather affects costs. The key assumption to make these instruments valid is that past weather does not affect demand.

2. Prices of intermediate goods (plants). Since I can trace the plants from the producer to the final sale, I can match all retail transactions with the sale price (from producer to processor) of the original plant. This instrument is valid if the producers either do not know or do not take into account the retail demand shocks. There are two reasons why this is a reasonable assumption: first, upstream transactions often happen months before the retail transactions and second, it is widely reported that producers in this market are facing essentially perfectly competitive conditions.

3. Wholesale prices in other markets. Because of the tracing system implemented in the state of Washington, I can match every product sold by a retailer to the wholesale transaction. This means that I observe the wholesale prices that retailers pay. Wholesale prices are determined by the wholesaler costs (supply shocks) and retailer
demand (demand shocks). If wholesalers have market power, wholesale prices will be correlated with $\xi_{jrt}$. This issue is particularly acute since different retailers may pay very different prices for the same product from the same wholesaler, in other words, wholesalers may be price discriminating. To avoid this problem, for each wholesaler-product-market-month, I construct the average wholesale price. Then for each retailer-wholesaler pair, I take the average over all markets in which the wholesaler is present, but the retailer is not. Finally, I take the average over all wholesalers that sell the given product to the retailer in the given month, weighted by the quantity. Note that different retailers within the same market will have different values of this instrument if they contract with different wholesalers. The biggest concern about the validity of this instrument is the risk that demand shocks are correlated across different markets. Note that common shocks across the entire state are controlled for by the time fixed effects.

My model has two non-linear parameters, the nesting parameter $\rho$ and the travel costs $\phi_1$ and $\phi_2$. I need sources of exogenous variation in conditional inside market shares to identify these two parameters (Berry and Haile (2014)). I use the following instruments: (i) number of other products in the market and (ii) the average of the cost-shifting instruments over all other products in the market. The intuition is that (i) if the number of other products increases, there is more competition and any given product is less likely to be picked; (ii) other instruments increase costs, which in turn increase prices for other products, which cause the share of the focal product to increase.

Table 2 presents the results of the first stage. All price instruments, except for temperature, are strongly correlated with prices.

2.4.2. Cost Estimation

I compute the marginal costs implied by equation (2.4). Note that the FOC is a linear system of equations with respect to marginal costs, so I can solve for them exactly. Denote
the estimated marginal costs $\hat{c}_{jrt}$. Some of the estimated marginal costs are negative, in which case I replace them with 0.01.

I run a standard OLS regression of the implied marginal costs on the set of fixed effects: time, retailer, product type:

$$\log(\hat{c}_{jrt}) = \lambda_j + \omega_r + \psi_m + \nu_{jrt} \quad (2.16)$$

2.5. Estimation Results

The results of the demand estimation are summarized in Table 3. I estimate five models. Usable cannabis is the most popular product, which is reflected in the product fixed effects for all models. Not using instruments for price (column “logit no IV”), I underestimate the price elasticity by around 50%. The median price elasticity that I find for the full model is 1.56, which is around half of the elasticity found by Hollenbeck and Uetake (2019). Only 3% of estimated elasticities fall in the inelastic range (i.e. elasticity less than 1).

I compute demand elasticity with respect to distance, which can be interpreted as percentage change in demand as a response to 1% increase in transportation costs. I find an elasticity of 0.19, which is less than 1/8 of price elasticity. This means that in monetary terms, a 1% increase in distance is equivalent to 0.12 % increase in price. For the final specification, I

<table>
<thead>
<tr>
<th></th>
<th>prices</th>
</tr>
</thead>
<tbody>
<tr>
<td>rainfall</td>
<td>2.73</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>temperature</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td>wholesale prices</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
</tr>
<tr>
<td>upstream prices</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.76</td>
</tr>
<tr>
<td>N</td>
<td>33,257</td>
</tr>
</tbody>
</table>

Table 2: First stage results, standard errors in parentheses
<table>
<thead>
<tr>
<th></th>
<th>logit no IV</th>
<th>logit-IV</th>
<th>NL-IV</th>
<th>logit-IV-distance</th>
<th>NL-IV-distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solid edible $\beta_2$</td>
<td>-2.47</td>
<td>-2.30</td>
<td>-0.71</td>
<td>-2.30</td>
<td>-0.91</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.15)</td>
<td>(0.03)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>liquid edible $\beta_3$</td>
<td>-3.39</td>
<td>-3.09</td>
<td>-0.96</td>
<td>-3.10</td>
<td>-1.23</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.05)</td>
<td>(0.20)</td>
<td>(0.05)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>Vape $\beta_4$</td>
<td>-1.31</td>
<td>-0.89</td>
<td>-0.28</td>
<td>-0.90</td>
<td>-0.38</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Other $\beta_5$</td>
<td>-3.53</td>
<td>-3.39</td>
<td>-1.05</td>
<td>-3.40</td>
<td>-1.33</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
<td>(0.03)</td>
<td>(0.22)</td>
<td>(0.03)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Price $\alpha$</td>
<td>-0.05</td>
<td>-0.07</td>
<td>-0.02</td>
<td>-0.07</td>
<td>-0.03</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Distance $\phi_1$</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-0.07</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Other $\phi_2$</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-0.04</td>
<td>-8e-4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Nesting $\rho$</td>
<td>-</td>
<td>-</td>
<td>0.69</td>
<td>-</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.07)</td>
<td></td>
<td>(0.08)</td>
</tr>
<tr>
<td>Market FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Retailer FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Month FE</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Median price elasticity</td>
<td>1.05</td>
<td>1.56</td>
<td>1.49</td>
<td>1.33</td>
<td>1.57</td>
</tr>
<tr>
<td>Median distance elasticity</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.22</td>
<td>0.19</td>
</tr>
<tr>
<td>N</td>
<td>33,257</td>
<td>33,257</td>
<td>33,257</td>
<td>33,257</td>
<td>33,257</td>
</tr>
</tbody>
</table>

Table 3: Demand estimation results, standard errors in parentheses

compute the aggregate elasticity for the category as a whole of 0.96, which implies a high degree of substitution within the cannabis category. This means that there are not many close substitutes for legal cannabis available to the consumers, be it illegal drugs, including cannabis, or legal intoxicating substances like alcohol.

I estimate retailer fixed effects in the demand and cost equations, i.e. the retailer types, which are a crucial component of the counterfactual. I plot the types in Figure 8. There is much more heterogeneity in cost types than in demand types. Another thing to note is that there does not seem to be a trade-off between the cost and the retailer quality: the retailers with high demand types tend to have low cost types and vice versa. This implies that we can think of retailer types as essentially unidimensional, with higher efficiency leading to both lower cost and higher demand.
2.5.1. Demand Model Validation

The key challenge for the computation of the auction counterfactual is how well I can predict the market outcomes for unobserved market structures, i.e. what would the equilibrium prices and quantities be in a market where some of the observed retailers are absent and some of the unobserved retailers are present? I perform a demand validation exercise to evaluate how well the proposed demand system handles this challenge.

To perform this evaluation, I use the observed changes in market structure and see how the model predicts the changes in market outcomes after those changes. In the data, I observe 157 instances of retailer entry and exit, excluding the first entrants in a market. Treating the new entrants as if I do not observe their demand and cost types, I predict the expected prices and quantities after entry and then compare them to the observed outcomes.

Note that since I am using the demand parameters estimated on the full sample (i.e. there is no training and holdout samples), this is primarily a test of the model’s ability to predict a retailer’s demand and marginal costs without observing their type. When a retailer’s
type is unobserved, I draw from the set of types of other retailers in the same market. This means that the task is particularly hard for markets with a smaller set of retailers, in which a single highly successful or unsuccessful incumbent firm can drive the predictions for entrants.

As a metric of the quality of my predictions, I compute the prediction error as a share of the observed value. In this subsection, I focus on the prediction errors for quantity. The validation procedure details and results for prices can be found in Appendix A. Tables 4 provides summary statistics for the prediction error in quantity. The first column presents the summary statistics for the quantities and prices aggregated to individual retailer level, in the second and third columns we can see the entrants and incumbents separately. The fourth column contains the summary statistics for market-level variables. Finally, in the fifth column I consider only markets with 6 firms and above. On the market level, especially for large markets, the model performs reasonably well. It is particularly encouraging that the mean and median error are close to zero. The individual predictions for entrants are severely overestimated. The model predicts a wide range of values, but market shares are already quite small and cannot go below zero, so the large errors cannot be matched by small ones.

As I am primarily interested in the results for the market overall and larger markets are more important, I conclude that my demand model performs reasonably well for the task at hand.
2.6. Counterfactual: Using Auctions Instead of Lotteries

2.6.1. Computation

There are two main computational challenges in computing the counterfactual. First, there are retailers whose types I do not observe. To overcome this problem, I draw the retailer types with replacement from the estimated distribution of retailer types conditional on being in the given county. I draw the entire vector of types for the unobserved retailers 500 times for each county and compute the lottery and auction outcomes for these draws, after which I take the average across them to get the expected auction and lottery outcomes.

The second challenge is that to perform the equilibrium selection procedure described in subsection 2.3.5, I need to take the expectation \( \mathbb{E}(\pi_r(\Omega)) \) over possible allocations \( \Omega \) repeatedly. The set of all possible allocations grows very quickly with both the number of applicants and the number of licenses. Again, I compute the expectation by taking Monte-Carlo draws from the underlying set of allocations and computing the profits for those draws. At each stage of the equilibrium selection, I perform a 1000 draws, until the full set of possible allocations is smaller than 1000. Note that this implies that at the last stage (i.e. when I am selecting the last retailer to receive a license, I always compute the full set of possible allocations and guarantee that the outcome I find is a Nash equilibrium).

Since the licenses are capped at the local level (i.e. at the town or city level or county at large), quite often there would be multiple auctions within a county. In this case, the outcomes in one auction affect the profits of retailers in the other auction. To deal with this issue, I run the auctions “simultaneously”, i.e. the first license in both auctions is allocated based on unconditional expectation, but the expectation for the second round of the auctions is computed conditional on both winners being in the market etc.

Once both the types and the set of retailers are drawn, I can compute the market equilibrium by iterating over the FOCs.
2.6.2. Revenue and market outcomes under auction and lottery

Table 5 summarizes the key market outcomes for the counterfactual auction and lottery. The first two columns show the expected annual quantity sold in kilograms. In all markets, the auction leads to an increase in quantity, with the overall increase of 5%. This is a relatively modest increase, which corresponds to the finding of high degree of substitution between products within the same market (i.e. high $\rho$). If a particular retailer becomes more efficient or lowers prices, most of their demand will come from other retailers rather than from expanding the market. Columns 4 and 5 show the quantity-weighted prices. Here we find that the auction leads to a 3% decrease in prices. As firms become more efficient, but can’t expand the market too much, the price competition becomes stronger. Finally, columns 6 and 7 show the expected yearly sales tax revenue (in $1000s) of the WA state government under lottery and auction respectively. I find that the tax revenue is 2% higher under the auction, as the decrease in prices is outstripped by the increase in quantity. Note that while the auction increases quantities in all markets, the effect for prices and tax revenue is mixed. The intuition is that higher quality, lower costs and more accessible locations all lead to an increase in quantity, while the effect on prices is mixed: higher quality and better locations increase prices, while lower costs lead to lower prices.

Table 6 summarizes more market outcomes for the auction and lottery. First, we look at the average demand ($\bar{\nu}$) and cost ($\bar{\omega}$) types for the market participants under the two allocation mechanisms. In case of lottery, these are just averages across the types observed in the given county. In case of auction, this is the average across the types that win the auction. The demand type is uniformly higher across all counties, while the cost type is typically lower with the exception of a few markets. This implies that auctions are mostly driven by selection on demand type, i.e. quality. The average increase in demand type across all markets is 29% and the average decrease in the cost type across all markets is 16%. In columns “Rmax” and “Rmin” I present the upper and lower bounds on the auction revenue in each market (in $1000s). The total auction revenue is estimated to be between $
Table 5: Lottery and auction outcomes by county: average quantity per year in kg, average weighted price, average annual tax income in $1000
The last two columns present the number of licenses that are auctioned (i.e., excluding the markets in which licenses were not oversubscribed) and the number of applicants for those licenses. As expected, the more oversubscribed markets see a higher auction revenue per license (e.g., compare Spokane county with 3 applicants per license to similarly-sized Clark and Pierce county with 9 and 5 applicants per license).

What is the price of fairness? Considering only the state’s revenue, the government receives an extra $8M every year in perpetuity in taxes from the auction in addition to the immediate auction revenue of at least $57.3M. Assuming a discount rate of 1% and taking the time horizon to be 10 years, the price of fairness is approximately $137M. For perspective, the WA state budget for 2014 was $35.1B, so the price of fairness is 0.39% of the state’s annual budget.

2.6.3. Racial disparities in gains from auction

I have established that by all aggregate metrics (tax revenue and consumer surplus), residents of the state of Washington would be better off if the retail cannabis licenses were allocated by an auction instead of lottery. However, policymakers may consider the distribution of efficiency gains from an auction, namely whether the demographic groups that were most negatively affected by drug criminalization are not left behind by the reform. In this subsection, I consider (i) the ownership of licenses under auction and lottery and (ii) the distribution of consumer gains from an auction. I establish that auction results in expected loss of 10 licenses for Black applicants, which is a 21% reduction. All of these licenses go to white applicants. The gains in consumer surplus (i) accrue disproportionately less to majority black and majority Asian areas; (ii) accrue disproportionately more to majority-white areas, and (iii) are positively correlated with the population share of white residents on the block group level. This disparity cannot be explained by differences in income, as lower income areas benefit more from the auction than higher income areas.

I start with ownership. Table 7 summarizes the results. Overall, the state population in
<table>
<thead>
<tr>
<th>County</th>
<th>$\bar{\nu}$ lottery</th>
<th>$\bar{\nu}$ auction</th>
<th>$% \Delta \bar{\nu}$</th>
<th>$\bar{\omega}$ lottery</th>
<th>$\bar{\omega}$ auction</th>
<th>$% \Delta \bar{\omega}$</th>
<th>$R_{\text{max}}$</th>
<th>$R_{\text{min}}$</th>
<th>$L$</th>
<th>$N$</th>
</tr>
</thead>
<tbody>
<tr>
<td>King</td>
<td>-0.51</td>
<td>-0.50</td>
<td>1%</td>
<td>0.19</td>
<td>0.19</td>
<td>1%</td>
<td>25,274.06</td>
<td>23,687.74</td>
<td>84</td>
<td>458</td>
</tr>
<tr>
<td>Pierce</td>
<td>-0.48</td>
<td>-0.46</td>
<td>5%</td>
<td>-0.95</td>
<td>-1.00</td>
<td>-5%</td>
<td>2,889.02</td>
<td>2,781.56</td>
<td>28</td>
<td>152</td>
</tr>
<tr>
<td>Snohomish</td>
<td>-0.36</td>
<td>-0.34</td>
<td>6%</td>
<td>-1.06</td>
<td>-1.71</td>
<td>-3%</td>
<td>5,133.50</td>
<td>4,934.16</td>
<td>38</td>
<td>149</td>
</tr>
<tr>
<td>Spokane</td>
<td>0.08</td>
<td>0.10</td>
<td>32%</td>
<td>-2.20</td>
<td>-2.19</td>
<td>1%</td>
<td>1,419.78</td>
<td>1,392.30</td>
<td>31</td>
<td>106</td>
</tr>
<tr>
<td>Clark</td>
<td>-0.01</td>
<td>0.01</td>
<td>195%</td>
<td>0.24</td>
<td>0.13</td>
<td>-48%</td>
<td>4,003.96</td>
<td>3,856.34</td>
<td>13</td>
<td>113</td>
</tr>
<tr>
<td>Thurston</td>
<td>0.02</td>
<td>0.05</td>
<td>156%</td>
<td>-2.98</td>
<td>-3.04</td>
<td>-2%</td>
<td>3,824.92</td>
<td>3,321.28</td>
<td>21</td>
<td>67</td>
</tr>
<tr>
<td>Whatcom</td>
<td>-0.03</td>
<td>0.00</td>
<td>98%</td>
<td>-1.41</td>
<td>-1.52</td>
<td>-7%</td>
<td>5,265.10</td>
<td>3,960.68</td>
<td>24</td>
<td>61</td>
</tr>
<tr>
<td>Kitsap</td>
<td>-0.44</td>
<td>-0.41</td>
<td>7%</td>
<td>-0.35</td>
<td>-0.36</td>
<td>-3%</td>
<td>2,404.33</td>
<td>2,160.75</td>
<td>19</td>
<td>76</td>
</tr>
<tr>
<td>Cowitz</td>
<td>0.29</td>
<td>0.34</td>
<td>17%</td>
<td>-2.53</td>
<td>-2.88</td>
<td>-14%</td>
<td>64.80</td>
<td>56.20</td>
<td>8</td>
<td>22</td>
</tr>
<tr>
<td>Skagit</td>
<td>0.12</td>
<td>0.21</td>
<td>71%</td>
<td>-0.92</td>
<td>-1.04</td>
<td>-14%</td>
<td>1,287.80</td>
<td>1,212.25</td>
<td>13</td>
<td>34</td>
</tr>
<tr>
<td>Benton</td>
<td>0.32</td>
<td>0.38</td>
<td>18%</td>
<td>-3.78</td>
<td>-4.13</td>
<td>-9%</td>
<td>890.05</td>
<td>785.19</td>
<td>3</td>
<td>37</td>
</tr>
<tr>
<td>Grays Harbor</td>
<td>0.05</td>
<td>0.16</td>
<td>226%</td>
<td>-4.70</td>
<td>-5.52</td>
<td>-17%</td>
<td>4,203.99</td>
<td>4,068.92</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>Yakima</td>
<td>0.32</td>
<td>0.34</td>
<td>9%</td>
<td>-4.60</td>
<td>-5.01</td>
<td>-9%</td>
<td>965.69</td>
<td>875.06</td>
<td>8</td>
<td>37</td>
</tr>
<tr>
<td>Clallam</td>
<td>-1.00</td>
<td>-0.88</td>
<td>13%</td>
<td>-0.90</td>
<td>-1.47</td>
<td>-64%</td>
<td>231.82</td>
<td>222.66</td>
<td>8</td>
<td>26</td>
</tr>
<tr>
<td>Island</td>
<td>0.14</td>
<td>0.32</td>
<td>128%</td>
<td>-2.32</td>
<td>-3.29</td>
<td>-42%</td>
<td>378.54</td>
<td>344.36</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Whitman</td>
<td>0.96</td>
<td>0.99</td>
<td>3%</td>
<td>-4.77</td>
<td>-5.47</td>
<td>-15%</td>
<td>73.71</td>
<td>66.03</td>
<td>4</td>
<td>9</td>
</tr>
<tr>
<td>Walla Walla</td>
<td>-0.07</td>
<td>0.01</td>
<td>116%</td>
<td>-1.43</td>
<td>-1.71</td>
<td>-20%</td>
<td>399.06</td>
<td>299.32</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Asotin</td>
<td>0.46</td>
<td>0.49</td>
<td>5%</td>
<td>-7.27</td>
<td>-11.38</td>
<td>-56%</td>
<td>431.76</td>
<td>334.95</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Chelan</td>
<td>1.55</td>
<td>1.63</td>
<td>6%</td>
<td>-7.84</td>
<td>-10.53</td>
<td>-34%</td>
<td>356.60</td>
<td>255.62</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Kittitas</td>
<td>0.41</td>
<td>0.49</td>
<td>20%</td>
<td>-2.69</td>
<td>-3.67</td>
<td>-36%</td>
<td>342.66</td>
<td>309.22</td>
<td>6</td>
<td>20</td>
</tr>
<tr>
<td>Mason</td>
<td>0.05</td>
<td>0.13</td>
<td>132%</td>
<td>-2.69</td>
<td>-3.05</td>
<td>-13%</td>
<td>600.79</td>
<td>553.55</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>Jefferson</td>
<td>0.57</td>
<td>0.67</td>
<td>16%</td>
<td>-2.00</td>
<td>-2.65</td>
<td>-32%</td>
<td>190.94</td>
<td>177.26</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>Grant</td>
<td>-0.28</td>
<td>-0.22</td>
<td>20%</td>
<td>-1.84</td>
<td>-2.26</td>
<td>-23%</td>
<td>443.87</td>
<td>346.40</td>
<td>8</td>
<td>21</td>
</tr>
<tr>
<td>Douglas</td>
<td>-0.31</td>
<td>-0.26</td>
<td>15%</td>
<td>-7.28</td>
<td>-8.49</td>
<td>-17%</td>
<td>79.28</td>
<td>(31.70)</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>Lewis</td>
<td>0.86</td>
<td>1.04</td>
<td>21%</td>
<td>-5.71</td>
<td>-5.08</td>
<td>11%</td>
<td>170.05</td>
<td>157.45</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>Klickitat</td>
<td>-0.11</td>
<td>0.28</td>
<td>289%</td>
<td>-4.63</td>
<td>-8.39</td>
<td>-81%</td>
<td>418.56</td>
<td>341.60</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Stevens</td>
<td>-0.45</td>
<td>-0.26</td>
<td>44%</td>
<td>-4.49</td>
<td>-5.51</td>
<td>-23%</td>
<td>56.26</td>
<td>49.65</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>Pacific</td>
<td>-0.35</td>
<td>-0.22</td>
<td>37%</td>
<td>-2.21</td>
<td>-3.39</td>
<td>-53%</td>
<td>341.11</td>
<td>256.28</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Okanogan</td>
<td>0.35</td>
<td>0.43</td>
<td>20%</td>
<td>-6.66</td>
<td>-7.71</td>
<td>-16%</td>
<td>77.01</td>
<td>71.90</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>San Juan</td>
<td>1.14</td>
<td>1.14</td>
<td>0%</td>
<td>-17.73</td>
<td>-17.73</td>
<td>0%</td>
<td>46.34</td>
<td>41.76</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>Skamania</td>
<td>-0.23</td>
<td>0.03</td>
<td>115%</td>
<td>-6.88</td>
<td>-11.77</td>
<td>-71%</td>
<td>143.42</td>
<td>113.55</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Ferry</td>
<td>0.81</td>
<td>0.81</td>
<td>0%</td>
<td>-22.66</td>
<td>-22.66</td>
<td>0%</td>
<td>151.64</td>
<td>53.16</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Adams</td>
<td>-0.24</td>
<td>-0.16</td>
<td>36%</td>
<td>-4.07</td>
<td>-6.88</td>
<td>-69%</td>
<td>20.34</td>
<td>18.70</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Pend Oreille</td>
<td>-1.15</td>
<td>-1.15</td>
<td>0%</td>
<td>-16.13</td>
<td>-16.13</td>
<td>0%</td>
<td>31.16</td>
<td>30.89</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Franklin</td>
<td>-1.14</td>
<td>-1.14</td>
<td>0%</td>
<td>-16.42</td>
<td>-16.42</td>
<td>0%</td>
<td>1.11</td>
<td>0.79</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

| Total        | -0.15                 | -0.11                 | 29%                      | -1.93                    | -2.23                    | -16%                     | 63,147.38       | 57,338.84       | 382 | 1583 |

Table 6: Lottery and auction results: average demand type ($\bar{\nu}$), average cost type ($\bar{\omega}$), maximum and minimum auction revenue, number of licenses allocated and number of applicants.
2010 was 77% white, 7% Asian, 3.6% Black, 1.5% Native American, 0.6% Hawaiian or Pacific Islander and 10% other or biracial. The applications columns suggests that Black residents of WA were disproportionately more likely to apply for a cannabis license. Under lottery, on average 38 of them would receive a license. However, under auction, only 30 would, which is a 21% decrease in the number of Black cannabis licensees under auction. Similarly, auction would lead on average to a loss of 1 license (10% of all licenses) by a Hispanic applicant. This indicates that the concerns about fair distribution of licenses are well-founded.

Now consider the distribution of consumer gains. Since the state does not collect data on individual consumers, I use geographic variation to identify racial disparities. Using 2010 census data, I collect the demographic compositions on block group level and combine it with my counterfactual estimates of consumer surplus. First way I try to proxy for racial groups is by identifying block groups in which that group counts for more than 50% of the population. Table 8 presents the results. While non-Hispanic and Hispanic white-majority areas enjoy a 20% and 27% increase in consumer surplus respectively, Black and Asian-majority areas see only a 3% increase. Native American-majority areas and areas with no demographic majority also receive less than the average increase, at 13% and 12% respectively.

Note that in absolute terms, according to my metric, every demographic group benefits from the auction. However, in relative terms, the benefits are disproportionately received by white WA residents. However, I am drawing conclusions from a relatively small number

Table 7: Predicted race of license owners under auction and lottery

<table>
<thead>
<tr>
<th></th>
<th>Applicants</th>
<th>Share</th>
<th>Lottery</th>
<th>Share</th>
<th>Auction</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-hispanic white</td>
<td>867</td>
<td>54.8%</td>
<td>211</td>
<td>55.4%</td>
<td>221</td>
<td>57.8%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>45</td>
<td>2.8%</td>
<td>11</td>
<td>2.8%</td>
<td>10</td>
<td>2.7%</td>
</tr>
<tr>
<td>Black</td>
<td>138</td>
<td>8.7%</td>
<td>38</td>
<td>10.0%</td>
<td>30</td>
<td>7.9%</td>
</tr>
<tr>
<td>Asian</td>
<td>120</td>
<td>7.6%</td>
<td>28</td>
<td>7.4%</td>
<td>28</td>
<td>7.4%</td>
</tr>
<tr>
<td>Name NA</td>
<td>413</td>
<td>26.1%</td>
<td>93</td>
<td>24.3%</td>
<td>92</td>
<td>24.2%</td>
</tr>
<tr>
<td>Total</td>
<td>1583</td>
<td>100%</td>
<td>382</td>
<td>100%</td>
<td>382</td>
<td>100%</td>
</tr>
</tbody>
</table>
of block groups, in particular, there is only 10 block groups that are majority Black and most Black people in Washington live outside of those block groups. To alleviate this problem, I divide the block groups by decile of share of white residents, i.e. group 0 are the blocks with the lowest share of white residents (5%-56%) and group 9 are the blocks with the highest share of white residents (97%-100%). The results are presented in table 9. Now all groups are approximately the same size. I find that the gains from the auction are increasing in the share of white residents. In particular, group 0 sees an 11% increase in consumer surplus and group 9 sees a 32% increase. Another common concern is that the market underprovides services in minority or poor areas (e.g. food deserts). For each block group, I compute the probability that at least one store is located within it. The last two columns in Table 9 show the probabilities of at least one store being located within an average block group of a given decile. There seems to be no relationship between the share of white residents and store location both under auction and under lottery.

Another important variable that is correlated with racial composition of a block group is income. If retailers receive higher profits in higher income areas, the stores in those areas will be more likely to win the auction, which would in turn explain the racial disparity documented above. To investigate this mechanism, I divide the block groups into deciles by median household income from 2010 census and compute the average consumer surplus gain from the auction for each decile. Figure 9 shows that in fact the highest income block groups tend to receive less benefit from the auction. Specifically, the lowest income areas

<table>
<thead>
<tr>
<th>Race</th>
<th>number of block groups</th>
<th>population</th>
<th>ΔCS</th>
<th>HH income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Hispanic white no majority</td>
<td>4,196</td>
<td>5,878,871</td>
<td>20%</td>
<td>65,423.10</td>
</tr>
<tr>
<td>Asian</td>
<td>455</td>
<td>656,232</td>
<td>12%</td>
<td>46,763.59</td>
</tr>
<tr>
<td>Hispanic white</td>
<td>31</td>
<td>56,105</td>
<td>27%</td>
<td>34,454.56</td>
</tr>
<tr>
<td>Native American</td>
<td>14</td>
<td>16,898</td>
<td>13%</td>
<td>36,765.28</td>
</tr>
<tr>
<td>Other</td>
<td>31</td>
<td>60,888</td>
<td>55%</td>
<td>31,940.00</td>
</tr>
<tr>
<td>Black</td>
<td>10</td>
<td>14,056</td>
<td>3%</td>
<td>31,746.87</td>
</tr>
<tr>
<td>Total</td>
<td>4,766</td>
<td>6,724,540</td>
<td>19%</td>
<td>60,392.86</td>
</tr>
</tbody>
</table>

Table 8: Consumer benefits of auction on the block level, divided by the race of majority of residents.
Table 9: Consumer benefits of auction on the block level, divided by decile of share of white residents

<table>
<thead>
<tr>
<th>Decile</th>
<th>Interval</th>
<th># of BGs</th>
<th>Population</th>
<th>HH Income</th>
<th>ΔCS</th>
<th>P(store) auction</th>
<th>P(store) lottery</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>[5%,56%]</td>
<td>477</td>
<td>700,354</td>
<td>56314.40</td>
<td>11%</td>
<td>5%</td>
<td>5%</td>
</tr>
<tr>
<td>1</td>
<td>[56%,68%]</td>
<td>476</td>
<td>723,748</td>
<td>59755.21</td>
<td>13%</td>
<td>8%</td>
<td>7%</td>
</tr>
<tr>
<td>2</td>
<td>[68%,75%]</td>
<td>476</td>
<td>684,491</td>
<td>62297.13</td>
<td>20%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>3</td>
<td>[75%,81%]</td>
<td>476</td>
<td>692,128</td>
<td>62332.02</td>
<td>15%</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>4</td>
<td>[81%,85%]</td>
<td>477</td>
<td>683,560</td>
<td>66208.13</td>
<td>20%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>5</td>
<td>[85%,88%]</td>
<td>476</td>
<td>655,089</td>
<td>65067.42</td>
<td>15%</td>
<td>6%</td>
<td>7%</td>
</tr>
<tr>
<td>6</td>
<td>[88%,91%]</td>
<td>476</td>
<td>660,942</td>
<td>68240.89</td>
<td>21%</td>
<td>7%</td>
<td>7%</td>
</tr>
<tr>
<td>7</td>
<td>[91%,94%]</td>
<td>476</td>
<td>680,314</td>
<td>63975.55</td>
<td>19%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>8</td>
<td>[94%,97%]</td>
<td>476</td>
<td>667,461</td>
<td>66793.58</td>
<td>21%</td>
<td>6%</td>
<td>6%</td>
</tr>
<tr>
<td>9</td>
<td>[97%,100%]</td>
<td>477</td>
<td>575,468</td>
<td>65207.88</td>
<td>32%</td>
<td>5%</td>
<td>5%</td>
</tr>
</tbody>
</table>

increase their consumer surplus by 21%, while the highest income areas only by 10%. This means that the racial disparity cannot be driven by income differences between demographic groups. Once again, we consider the location of the retailers. Figure 10 shows the average probability that a store is located within a block group by income decile. Stores are less likely to be located in lower-income areas. The allocation mechanism has little to no effect.

2.6.4. How much fairness is lost?

I started this paper from the premise that lottery is chosen by the WA state government as an allocation mechanism for cannabis licenses because of fairness considerations. In this section I try to quantify the difference in fairness between a lottery and an auction in order to test my assumption.

A lottery gives all participants within a market an equal chance of winning, so I compute the probability of winning in lottery and in auction conditional on (i) location and (ii) type. Figures 11 and 12 show these probabilities, sorted from highest to lowest. Note that the order of retailers is not the same for lottery and auction, so the same position on the x axis does not necessarily correspond to the same retailer.

First, note that the probability of winning is far from uniform in a lottery, as some markets are more oversubscribed than others. Second, in both figures there is a similar pattern of
Figure 9: Percentage increase in consumer surplus from the auction as a function of median household income decile.

Figure 10: Probability of at least one store being present in an average block group as a function of median household income decile.
Figure 11: Probability of winning conditional on location, from the least to the most successful location, for auction and lottery

Figure 12: Probability of winning conditional on type, from the least to the most successful type, for auction and lottery
increasing probability on the upper end of the distribution and of decreasing on the lower end. Quantitatively, the top 20% locations have an average 52% chance of winning under auction and only 43% under lottery, while the bottom 20% of locations have an average 7.5% chance of winning under auction and 11% chance of winning under lottery. The results are even more dramatic for types: for the top 20% of types, the chance of winning on average is 61% under auction and 48% under lottery; for the bottom 20% of types, the chance of winning is 11% under auction and 16% under lottery.

Does the increase in “inequality of opportunity” under auction spill over to consumers? One possible concern is that under auction, all retailers will move to densely-populated, urban areas and rural consumers will be worse off. Figure 13 shows the map of which areas in Washington state are better off under auction (in yellow), and which are better off under lottery (in purple). It is clear that many of the states rural areas are better off under auction, while some urban areas—most notably Spokane—are worse off. The correlation between population density and percentage difference between the consumer surplus under auction and lottery is 0.1, so the relationship is at best very weak. I also use the list of census designated places (CDP) in WA to identify urban and rural blocks and compute average population-weighted difference between consumer surplus under auction and lottery. I define a block as urban if it lies within a CDP with ”City” in its name and rural if it either lies in a CDP that is neither city nor town or if it is not a part of any CDP. I find a 20% increase in consumer surplus in urban areas and 25% increase in consumer surplus in rural areas. I conclude that rural areas would not be on average worse off under an auction.

I also consider the overall spatial inequality, beyond the urban/rural division. Does the auction increase inequality in consumer surplus between block groups? I compute the Gini coefficient for the distribution of consumer surplus across all block groups. I find that, under lottery, the coefficient is equal to 0.6, while under auction, it is 0.62, so there is 3% increase in spatial inequality. Figure 14 shows the Lorentz curve for the distribution of consumer surplus. The curve for a lottery is very slightly to the left of the curve for an auction.
Figure 13: Areas that are better off under auction (yellow) vs under lottery (purple)

Figure 14: Inequality in average consumer surplus, by block group, under auction and lottery
2.7. Conclusion

Many managers are now facing a choice between profit-maximization (a traditional goal of private enterprise) and fairness (due to increasing public pressure to incorporate it in the decision-making process). I study this trade-off in the context of cannabis legalization in the state of Washington. Questions of fairness are often raised when cannabis legalization is discussed. As marginalized groups bore the brunt of negative effects of drug criminalization and the war on drugs, it is only fair that members of those groups should be the major beneficiaries of legalization. However, the explicit goal of legalization was to raise tax revenue for the state budget. The policymakers chose to allocate retail licenses by lottery. My goal in this paper is to predict the license allocation if the state chose to allocate licenses by auction and compare the outcomes.

To achieve that goal, I estimate the equilibrium in the Washington cannabis market using the detailed data provided by the state Liquor and Cannabis Board (WLCB), which keeps track of every ounce of product from growing and harvesting to the final retail sale. I estimate the price elasticity, the travel cost for consumers, the marginal costs, and the distribution of heterogeneity among retailers in terms of their quality (i.e. their effect on demand) and cost effectiveness. I then use these estimates to simulate counterfactual license allocations under lottery and auction mechanisms. The key challenge is predicting the behavior of retailers who applied for a license, but did not win, and therefore are not observed in the data. I use the application data from WLCB for their location information and draw their types from the estimated distribution of heterogeneity.

The results show that (i) the auction does in fact select more efficient retailers (ii) it leads to a market expansion of 5% and (iii) a 3% reduction in prices. The auction leads to a 2% or $8M increase in state tax revenue compared to lottery. Taking into account the auction revenue, the price of fairness over a 10 year window is $137M. There is no tradeoff between state revenue and consumer surplus (typical for “sin goods”): under the auction, cannabis consumers are substantially better off: they visit higher quality retailers and pay lower
Using a machine learning algorithm to predict the applicants’ racial identity from their first and last name, I find the expected number of licenses going to each demographic group under auction and lottery. I find that under auction Black applicants receive 21% less licenses and Hispanic applicants receive 10% less licenses.

I document racial disparities in the distribution of consumer benefits from the auction. First, while majority-white block groups receive 20% higher consumer surplus under auction, in majority-Black areas consumer surplus increases on average by only 3%. In general, the consumer gain from auction is increasing in the population share of white residents: while the least white block groups receive 11% higher consumer surplus, the most white block groups receive 32% higher consumer surplus. This disparity cannot be explained by differences in median household income, as the increase in consumer surplus is decreasing in median household income, so controlling for income, the racial disparity should be even stronger. These findings imply that concerns over fairness were well-founded.

Finally, I try to quantify the amount of equity lost when using an auction instead of a lottery. I compute the a priori probability of winning a license conditional on a firm’s location or type under both mechanisms. The lottery does not provide perfectly equal chances of winning to everybody, as different markets have different numbers of applicants for a different number of licenses. However, under an auction, we see a substantial increase in “inequality of opportunity” conditional on location and an especially high increase in inequality conditional on retailer type. Another concern is spatial inequality among consumers: does an auction lead to all retailers concentrating in urban centers leaving the rural areas worse off? My model allows for the investigation of this question by estimating consumer surplus on 2010 census block group level. I find that the consumer gains or losses are not correlated with population density or whether the area is urban or rural. In fact, I find that while urban areas are on average 20% better off under auction, rural areas are 25% better off. I also find that the Gini coefficient of (population-weighted) spatial inequality is 3% higher under
auction compared to lottery, so spatial inequality is only slightly higher under auction.

My findings could be useful for platform companies who increasingly have to incorporate fairness considerations into the platform design. I show that a switch from a mechanism that prioritizes efficiency, such as auction, to a mechanism that prioritizes equity, such as a lottery, can lead to lower profits for the platform and lower consumer surplus. However, practitioners should keep in mind that the benefits of the increased efficiency may be unequally distributed, raising fairness considerations.
3.1. Introduction

An increasing number of markets for products and services is intermediated through two-sided platforms. Often platforms also produce some of the products that compete on these markets, raising concerns from policymakers and the press about preferential treatment the platforms may show towards their own products. For example, competition authorities have raised concerns that Google uses its search engine to promote own products (e.g. Pixel phones, Nest smart thermostats). Similarly, Amazon features its own private label products prominently in the search results for most product categories. Defenders of these companies point out that such behavior does not necessarily imply bias, since those products may be the ones that meet customer needs best. These settings highlight a well-known tension between pro- and anti-competitive aspects from vertical integration. On the one hand, the integrated firm may increase rivals costs or make it harder to reach consumers, leading to potential foreclosure. On the other hand, the elimination of double marginalization creates efficiencies from integration. In these newer settings, the ‘input’ provided by the platform is preferential access to advertising exposure, which may give competitive advantage to the integrated product.

The market for television advertising presents a similar setting, as three out of the four major television broadcasters are owned by a media conglomerate that also owns a major movie studio: in 2011 the relationships are NBC and Universal, ABC and Disney, and FOX and 20th Century. Television advertising is an important competitive tool for movie studios, as promotional expenditures typically account for about one third of the full cost of a movie (from Elberse and Anand (2007), for discussion of role of advertising in competition between

---

1 Based on work with Sylvia Hristakeva and Julie Mortimer
2 The main broadcasters in the U.S. are ABC, CBS, NBC, and FOX. Ten years later, in 2021, it is again three out of four, but in a different configuration, with CBS owned by Viacom and FOX separated from 20th Century.
studios, see also Eliashberg et al. (2006), Rao et al. (2017)). In addition, industry practitioners agree that integrated studios benefit from preferential treatment when advertising on sister networks. That is, affiliation between a movie studio and a broadcast network is likely to give it access to cheaper advertising; therefore, one expects that affiliated networks would benefit from a competitive advantage.

We also see the preferential treatment of integrated services in the times of “streaming wars.” Companies that are launching own streaming services are extensively relying on own platforms to reach customers. For example, NBCUniversal planned to spend more than twice as much on own television channels and platforms to promote its Peacock service; Disney+ and Paramount+ also rely on the promotional ad inventory of the integrated company.\(^3\) Furthermore, in 2019 Disney banned Netflix ads on its television networks. Initially, Disney “put out an edict to staffers that it wouldn’t accept ads from any rival streaming services, but later reversed course and found a compromise with nearly every company, the people familiar with the situation said. The exception was Netflix.”\(^4\)

Our empirical analysis investigates whether movie studios benefit from their affiliation with broadcast television networks. An ideal dataset would include not only product-market data, but also information on advertising quantities and individual prices paid by movie studios at each broadcast network. Unfortunately, advertising prices are considered trade secrets and are rarely available to researchers. Instead, our approach focuses on movie studios’ advertising exposure across broadcast networks. We ask: do integrated studios direct a higher share of advertising exposure to their affiliated networks? The analysis is based on the Rentrak dataset from 2011 -2013, which tracks all instances of national advertising with a rich set of characteristics, including time, position during within the ad break, copy, product, brand, advertiser and parent company, as well as telecast characteristics, such as viewership among various demographics. Information on advertising exposures across

---


76
products is more easily accessed by researchers, say by scraping websites, than getting data on advertising prices or product market sales; hence, our approach may be applied to many settings where researchers do not have detailed or proprietary product market information.

A major challenge for this analysis is to separate the affiliation incentives from potential unobserved match values between the affiliated studio and the network. Suppose that we find that Disney advertises more on ABC than on other broadcast networks. This may be explained by a vertical integration rationale in which ABC takes into account Disney’s profits and sets a lower advertising price for the movie studio. However, the observed exposure may also be explained by an unobserved match between Disney’s movies and ABC’s programming. For example, if ABC’s programming is likely to reach the audience of interest for the average Disney movie.

To separate these two effects, we use a matching strategy in the spirit of DellaVigna and Hermle (2017), which exploits the advertising exposures of all products to capture the unobserved “preferences” that advertisers have for networks’ content. We exploit the detailed information available in the ad-placement dataset, where we observe all national ad placements for more than 30,000 distinct products. For each movie we identify the products with similar advertising strategies and hence similar target audiences. To do so, we use clustering based on product advertising exposure on non-affiliated cable networks, such as FOOD, AMC, BET, which likely reach different audiences. This approach allows us to control for potential unobserved match values between a movie and the programming on broadcast. We find a lot of variation in the products matched to different movie types.

To investigate efficiencies from vertical integration we check whether integrated studios advertise relatively more on affiliated networks than other products with similar advertising strategies. We implement the analysis with a difference-in-differences approach, where the matched products serve as a control group. We find significant affiliation effects for ABC-Disney and FOX-20th Century relationships, but not for NBC-Universal. We also confirm the robustness of our findings by testing for heterogeneity in treatment effects across movie
budgets and genres. The intuition is that if the estimated treatment effect is driven by affiliation, then movie characteristics should be irrelevant, as they are already controlled for by the matched products. However, if some of the effect is driven by the movie content, then it would show up as heterogeneity, since movies of different genres target different audiences. We find no evidence of significant heterogeneity.

Our result suggests a wedge between affiliated and non-affiliated studios, which is likely driven by increases in efficiency. A simple theoretical example shows that ABC has incentives to both decrease the advertising cost to Disney and to raise the cost to a competitor, e.g. Paramount Pictures. To shed light on the potential anti-competitive effects from vertical integration, we investigate whether broadcast networks may put competing movie studios at a disadvantage. Here, we match unaffiliated studios to other products based on their advertising strategies on (unaffiliated) cable networks. Then, we ask whether, e.g. Paramount Pictures advertises less on ABC than other products with similar advertising strategies. We do not find significant effects.

Our findings are relevant to policymakers and practitioners considering the implications of vertical relations in the context of two-sided markets. We do not explicitly model the downstream market and quantify the efficiency effects. For many settings of policy interest, researchers do not have access to sales and pricing data, which would allow us to quantify the effects. As a result, our approach may be more widely applied to inform potential efficiency and anti-competitive effects across two-sided markets.

3.2. Theoretical motivations

The welfare effects of vertical integration are ambiguous because it may both increase efficiency for the integrated firm and have anti-competitive effects for the other players in the market. Vertical integration allows the firm to align incentives along the channel, which we typically see through a decrease in double marginalization. In the market for television advertising, the decrease in double marginalization implies that the affiliated studio
would likely access advertising inputs at lower prices and increase its advertising exposure. The anti-competitive effects come from the concern that the integrated television network has incentives to charge higher prices to rival studios, likely decreasing their advertising exposures and market shares.

We introduce a stylized theoretical example to showcase the efficiency and foreclosure concerns in the analyzed market. Suppose there are two movie studios, Disney (D) and Paramount (P) and one television network selling advertising exposures, ABC. Studios compete à la Hotelling with Disney located at 0 and Paramount located at 1. We set downstream prices to 1 for simplicity; this also matches with the industry practice that studios do not compete on prices of movie tickets. Instead, studios rely on advertising to increase demand. We model advertising in this market as persuasive, that is, advertising increases consumers’ utilities of the movie directly. The interpretation is that the social aspect of watching a movie accounts for a large part of the utility from watching a movie. Movies that are widely advertised are seen by a broader audience and create more discussion between the people who saw it and fear of missing out for those who did not. Let consumer $i$’s utility from going to see a movie by studio $j = \{D, P\}$ be:

$$U_{ij} = v + \alpha_j A_j - \gamma_j A_j^2 - d_{ij}. \tag{3.1}$$

We borrow the linear-quadratic specification from Chintagunta and Jain (1992) and Chintagunta (1993). $d_{ij}$ is the distance between consumer $i$ and studio $j$, capturing the taste match between the consumer and the studio. Advertising is persuasive, captured by $\alpha_j > 0$, but its effect is decreasing with the amount of advertising, captured by $\gamma_j > 0$. Note that

---

5We assume that ABC’s content is homogeneous. An alternative assumption is that we are considering the advertising market only for a specific TV show and different shows are not substitutes from the studios’ perspective, i.e. different shows are watched by non-overlapping groups of viewers. Note that, given the large share of promotional advertising in networks (ABC advertising other ABC programming), this market is unlikely to face capacity constraints.
this formulation implies that there is a utility-maximizing level of advertising:

\[ A_j^{SO} = \frac{\alpha_j}{2\gamma_j}. \]  

(3.2)

To highlight some of the empirical complications, we allow that studios are heterogeneous in advertising effectiveness. The pair \((\alpha_j, \gamma_j)\) may capture: (i) differences in the match value between a studio and the network’s programming and viewers; (ii) differences in the quality of each studio’s ad copy. Under the first interpretation, ABC’s content may attract viewers who are more likely to be affected by Disney’s advertising. The idea that Disney’s advertising for “Frozen” may be more effective during family programming than content attracting single men, and ABC may be creating content that is more family friendly than the other broadcasters. We focus on the first interpretation because it highlights the empirical complications of controlling for match value between a studio and its affiliated networks. In addition, we analyze major studios that invest heavily in advertising; thus, it is unlikely that any studio would have a persistent disadvantage in the quality of its ad copy (talented video editors/copywriters can be poached, and techniques can be copied).

Let \( v > 1 \), so consumers choose one of the two movies, and the consumer at \( \tilde{x} \) is indifferent between the two movies:

\[ \tilde{x} = \frac{1}{2}(1 + \alpha_D A_D - \gamma_D A_D^2 - \alpha_P A_P + \gamma_P A_P^2). \]  

(3.3)

All consumers in the \([0, \tilde{x}]\) interval prefer the Disney movie and all consumers in \([\tilde{x}, 1]\) prefer the movie by Paramount. Then assuming a uniform distribution of consumer tastes, the demand for Disney is \( D_D = \tilde{x} \), demand for Paramount is \( D_P = 1 - \tilde{x} \).

In this market, networks often charge different prices to their advertisers, so we allow that ABC may set a different price per unit of advertising for Disney and Paramount, \( p_D \) and \( p_P \). Profits for studio \( j \) is \( \pi_j = D_j - p_j A_j \). Maximizing with respect to \( A_j \) gives us the
following demand for advertising:

$$A_j^*(p_j) = \frac{\alpha_j - 2p_j}{2\gamma_j}$$

(3.4)

Comparing equations (3.2) and (3.4), we see that as long as \(p_j > 0\), the advertising level is below utility-maximizing level \(A_j^{SO}\). This occurs because ABC sets prices above marginal costs and charges a higher price to the studio with less elastic demand. With integration ABC would set prices taking into account Disney’s profits. Let \(\lambda\) track the level of internalization as in Crawford et al. (2018). Normalizing marginal cost of advertising to zero, the ABC’s payoff is:

$$V(p_D, p) = p_D A_D^*(p_D) + p_P A_P^*(p_P) + \lambda \pi_D(A_D^*(p_D), A_P^*(p_P)).$$

(3.5)

The payoff is maximized at the following prices:

$$p_D^* = \frac{\alpha_D(1 - \lambda)}{2(2 - \lambda)}$$

(3.6)

$$p_P^* = \frac{\alpha_P}{2(2 - \lambda)}.$$  

(3.7)

Plugging the prices into the demand functions, we get the equilibrium advertising levels:

$$A_D^* = \frac{\alpha_D}{2\gamma_D(2 - \lambda)}$$

(3.8)

$$A_P^* = \frac{\alpha_P(1 - \lambda)}{2\gamma_P(2 - \lambda)}.$$

(3.9)

We see that \(p_D^*\) is decreasing in \(\lambda\) and \(p_P^*\) is increasing in \(\lambda\), that is, as the studio and the network become more integrated, the network charges Disney a lower price and Paramount faces a higher price. The equilibrium advertising quantity is increasing in the match quality \(\alpha_j/\gamma_j\). It is increasing in the level of integration \(\lambda\) for the integrated firm and decreasing for its competitor.

81
Figure 15: Market share of Disney as a function of level of integration \( \lambda \), \( v = 1, \bar{u}_D = 2, \bar{u} = 3/2 \).

For example, if \( \lambda = 1 \), then \( p^*_D = 0, A^*_D = A^*_{SO_D} \): that is, with full internalization, ABC sets Disney’s price to marginal cost, eliminating double marginalization. In this market the efficiency gains imply an increase in advertising for Disney. However, integration also has an anti-competitive aspect, because ABC has incentives to charge higher prices for Paramount’s movie. Paramount’s advertising level gets further away from the utility-maximizing level as \( \lambda \) increases, and eventually goes to zero as advertising becomes prohibitively expensive. This stylized example suggests that a fully integrated firm forecloses the competing studio from the television advertising market, which is an important input for the “production of the final good.”

The empirical analysis investigates how vertical integration affects advertising quantity for the affiliated studio, e.g. Disney, and its competitors, e.g. Paramount. Looking closely at \( A^*_j \) illuminates the key empirical challenge: the need to separately identify the match value between the movie and the advertising input, \( \alpha_j \gamma_j \), and the integration effect, \( \lambda_j \). In section ?? we discuss our approach, which relies on matching movies to other products with similar advertising exposures. Our main identifying assumption is that the set of matched products has a similar \( \alpha_j \gamma_j \) as the analyzed movie. The analysis shows that Disney advertises relatively
more on ABC than on unaffiliated broadcast networks. As a result, we expect that Disney is gaining competitive advantage through integration. The theoretical example shows that differences in the access to the advertising market may affect product market competition. Figure 15 shoes that, in equilibrium, the market share and profit of Disney is increasing in $\lambda$. If integration has inflationary effects on rivals prices for advertising inputs, then market share and profit of Paramount would be decreasing in $\lambda$.

For completeness, we consider the welfare effects of integration. We assumed zero marginal costs of advertising and all payments between consumers, studios and the network cancel each other out; thus, total welfare is equal to consumers’ equilibrium utility.

\[
U^*(x) = \begin{cases} 
  v + \alpha_D A_D^* - \gamma_D (A_D^*)^2 - x, & \text{if } x < \bar{x} \\
  v + \alpha A^* - \gamma (A^*)^2 - (1 - x), & \text{if } x \geq \bar{x}
\end{cases}
\]  

(3.10)

For the purposes of this exercise, we leave the persuasive advertising as part of consumer welfare: advertising in the utility function proxies for the social benefits from seeing a film. Note that every part of the utility except the distance is common for all consumers. Denote the following:

\[
u_j^* = v + \alpha_j A_j^* - \gamma_j (A_j^*)^2
\]  

(3.11)

\[
\bar{u}_j = v + \alpha_j A_j^{SO} - \gamma_j (A_j^{SO})^2
\]  

(3.12)

which is the homogeneous part of utility in market equilibrium (3.11) and when the advertising is set to a socially optimal level (3.12). Plugging in the expressions (3.2), (3.8) and (3.9) and doing some algebra, we can show that:

\[
u_D^* = v + (\bar{u}_D - v) \frac{3 - 2\lambda}{(2 - \lambda)^2}
\]  

(3.13)

\[
u^* = v + (\bar{u} - v) \frac{(1 - \lambda)(3 - \lambda)}{(2 - \lambda)^2}
\]  

(3.14)
In these expressions, the match value is captured by the maximum utility \( \bar{u}_j \), which does not depend on \( \lambda \). Total welfare is equal to:

\[
W = \tilde{x}(u^*_D - \frac{\tilde{x}}{2}) + (1 - \tilde{x})(u^* - \frac{(1 - \tilde{x})}{2})
\] (3.15)

Plugging in \( \tilde{x} = \frac{1 + u^*_D - u^*}{2} \) and rearranging, we get the following:

\[
W = \frac{(u^*_D + 1)^2 + (u^* + 1)^2 - 3}{4}
\] (3.16)

We take the derivative using chain rule:

\[
\frac{\partial W}{\partial \lambda} = \frac{1}{(2 - \lambda)^3} \left( (u^*_D + 1)(1 - \lambda)(\bar{u}_D - \bar{v}) - (u^* + 1)(\bar{u} - \bar{v}) \right)
\] (3.17)

which is a continuous function of \( \lambda \) on \([0, 1]\). It is easy to see that at \( \lambda = 0 \), \( \frac{\partial W}{\partial \lambda} < 0 \). If \( \lambda = 1 \), the derivative reduces to \( (\bar{u}_D - \bar{u})(1 + \frac{3}{4}(\bar{u}_D + \bar{u}) - \frac{3}{4}) > 0 \) if \( \bar{u}_D > \bar{u} \). This means that there is a welfare-maximizing level of integration \( \lambda^* \), which lies in the interior of the interval \([0, 1]\), neither the full integration or the complete separation are socially optimal.

An example of the relationship between welfare and the level of integration is presented in Figure 16. The intuition is the following. Increasing \( \lambda \) increases Disney’s advertising and decreases Paramount’s advertising. When \( \lambda \) is small, the main effect comes from the fact that \( \bar{u}_D > \bar{u} \), i.e. advertising is shifted from a worse-matched product to a better-matched one. However, since the utility function is concave in the level of advertising, as \( \lambda \) increases, the marginal effect of Disney’s advertising diminishes, while the marginal effect of Paramount’s advertising goes up. When \( \lambda \) gets close to one, an increase in \( A_D \) leads to close to zero increase in utility, but a decrease in \( A \) has a stronger effect. Therefore, total welfare starts decreasing in \( \lambda \) for larger values of \( \lambda \). Note that the implication here is that vertical integration may lead to welfare loss even if the product of the integrated firm is better in a vertical sense.
3.3. Data and Background

The main data source for this chapter is the Rentrak corporation. The dataset covers a period of 3 years (January 2011 - December 2013) and contains television viewership data from 13 million households and 29 million cable set-top boxes. The viewership data is combined with extensive information on the advertisements aired during the show, including time, position within ad block, ad copy, product, product category, brand, advertiser and parent company. The total number of advertising spots in the data is more than 58 million, of which movie advertisements account for 1.36 million.

The primary focus of this project is on the relationships between major movie studios and broadcast TV networks. We focus on broadcast television specifically as it is far more popular than cable and attracts a much broader audience, thus more likely to have a significant effect on downstream competition between the studios. During the sample period, there are five major Hollywood studios: Universal Pictures, Walt Disney Pictures, 20th Century Fox Pictures, Paramount Pictures and Sony Pictures (formerly known as Columbia Pictures), all of which are owned by media conglomerates, as described in Table 10. There are 5 national broadcast TV networks: ABC, CBS, FOX, NBC and CW. However, during the
sample period CW (a joint venture of CBS and Warner Brothers) is a recent creation and relatively unpopular, compared to the other 4 broadcast networks, so we do not focus on CW in our analysis. Out of the remaining four broadcasters, 3 are affiliated with a major movie studio, the only exception being CBS, which was spun off from Viacom in 2006 and re-merged in 2019.

The focus of our paper are the relationships between NBC and Universal, ABC and Disney, and FOX and 20th Century. The three affiliated studio-network pairs have different histories, which potentially imply differing levels of integration, as it may take time for the conglomerate to adjust the management structure post-merger. The most recent of the three relationships is NBC-Universal, created in 2004 after Universal was sold to GE, which then owned NBC. In 2009, it was announced that Comcast was going to purchase a controlling stake in NBC-Universal. The deal was completed in 2011, after receiving conditional approval from the regulators. The oldest relationship is between 20th Century Fox and FOX, starting in 1986 at the inception of FOX broadcasting network. Both companies stayed under control of News Corp from 1986 and until the end of the sample period. ABC was purchased by Walt Disney Company in 1996.

All major film studios have “sister” minor studios: Focus Features for Universal, Touchstone for Walt Disney, Fox Searchlight and New Regency for 20th Century Fox. Theoretically, these studios should be facing the same affiliation effects as the majors. However, there are reasons to believe that these studios will exhibit a weaker affiliation effect compared to the major studios. First, these studios are producing more niche films, which are less likely to make a substantial profit, which reduces incentives for integration. Second, by their nature, these studios are meant to be “independent” from the more corporate major studios. In the rest of the paper, we will compare the results for the entire set of movies produced under the parent company to only the ones produced by the major studio.

We observe 196 movies by affiliated studios in our sample. Table 11 presents the number of movies in each genre produced by members of each conglomerate of interest. The genres
<table>
<thead>
<tr>
<th>Parent company</th>
<th>Major Studio</th>
<th>Broadcaster</th>
<th>Cable</th>
<th>Minor studios</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comcast Corp</td>
<td>Universal</td>
<td>NBC</td>
<td>Bravo, E!, Syfy, MSNBC, USA, CNBC &amp; others</td>
<td>Focus Features, Chiller Films</td>
</tr>
<tr>
<td>Walt Disney Co</td>
<td>Walt Disney</td>
<td>ABC</td>
<td>ABC Family, Disney XD, ESPN</td>
<td>Touchstone, Buena Vista</td>
</tr>
<tr>
<td>News Corp</td>
<td>20th Century Fox</td>
<td>FOX</td>
<td>Fox News, FX, FBN, National Geographic</td>
<td>Fox Searchlight, New Regency</td>
</tr>
<tr>
<td>Viacom</td>
<td>Paramount</td>
<td></td>
<td>BET, CMT, VH1, MTV, Nickelodeon, TV Land &amp; others</td>
<td>FilmDistrict</td>
</tr>
<tr>
<td>Sony</td>
<td>Sony (Columbia)</td>
<td></td>
<td>GSN</td>
<td>Screen Gems</td>
</tr>
<tr>
<td>Time Warner Inc</td>
<td>Warner Bros</td>
<td>CW</td>
<td>CNN, TBS, TNT, Cartoon, truTV, TMC, HLN</td>
<td>New Line, Picturehouse</td>
</tr>
<tr>
<td>CBS Corp</td>
<td></td>
<td>CBS, CW</td>
<td>TV Guide, CBS Sports</td>
<td>CBS Films</td>
</tr>
</tbody>
</table>

Table 10: Major television and movie studio ownership structure in 2011-2013
are taken from IMDB data. Comcast and News Corp produce substantially more movies than Disney in the sample period. No studio follows a strict specialization strategy, all three produce in all genres. Unsurprisingly, Disney specializes in animation and family films. Comcast/Universal concentrates on action and comedy, News Corp/20th Century Fox focuses on dramas. This suggests that average movies by these studios are different from each other and therefore likely to advertise differently, independent of their affiliations.

3.4. Analysis

We start by establishing whether affiliated studios do in fact place more advertising on affiliated broadcast networks. To account for differences in rating and ad duration, we construct ad exposure variable by multiplying the number of households watching a given show during which the ad airs by the duration of ad spot in seconds. In other words, advertising is measured in household seconds. It is important to take ratings and duration into account, as opposed to the raw number of ads because studios typically pay per thousand viewers on the basis of ad duration.

Given the advertising exposure, we can find the shares of each network in each parent company’s movie advertising portfolio. Table 12 presents the results. We can see that ABC is the top network for Disney and NBC is the top network for Comcast. FOX is not the top network for News Corp, however, FOX is relatively unpopular with movie studios in general and News Corp advertises on FOX disproportionately more than others. We can use the two unaffiliated studios, Time Warner and Sony, as a control group and compare the observed advertising allocation between affiliated and unaffiliated networks. If we use Warner as a

Table 11: Number of films in each genre by parent company

<table>
<thead>
<tr>
<th>parent</th>
<th>action</th>
<th>animation/family</th>
<th>comedy</th>
<th>drama</th>
<th>other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comcast Corp</td>
<td>23</td>
<td>6</td>
<td>27</td>
<td>16</td>
<td>6</td>
<td>78</td>
</tr>
<tr>
<td>News Corp</td>
<td>15</td>
<td>16</td>
<td>20</td>
<td>19</td>
<td>4</td>
<td>74</td>
</tr>
<tr>
<td>Walt Disney Co</td>
<td>11</td>
<td>23</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>44</td>
</tr>
</tbody>
</table>
control, only ABC-Disney relationship has a positive effect (13.3% vs 12.6%), since Warner allocates more advertising to NBC and FOX than Comcast and News Corp respectively. However, Warner’s advertising mix is more skewed towards broadcast television than any other studio. If we use Sony(Columbia) as a control group, all effects are positive. We get mixed results from this simple version of the analysis. It underscores that the choice of correct control group is crucial.

To perform our analysis more formally with a larger set of controls, we disaggregate to the level of individual movies. In the next section, we discuss the selection of the control group for each movie. Suppose that there is a number of control products that have the same advertising strategy as the focal movie, but are unaffiliated from the broadcast network of interest. For each of these products $i$, we compute the share of the focal network in their advertising allocation, across both the broadcast and cable TV channels. Then for each network (i.e., ABC, NBC, FOX), we run the following regression:

$$ s_i = \delta + \beta \times AffiliatedStudio_i + \varepsilon_i \quad (3.18) $$

where $AffiliatedStudio_i$ as a dummy for movie by the studio affiliated with the given network, $s_i$ is the share of TV advertising allocated to the given network by movie $i$. This is essentially the same as comparing the average share of, say, ABC, among the control products to the average share of ABC among Disney movies. If $\beta > 0$, this indicates that Disney advertises on ABC disproportionately.

Note that the specification (3.18) does not take into account the differences in how much

---

**Table 12: Major studios’ TV advertising allocation, affiliated studio-network pairs in bold.**

<table>
<thead>
<tr>
<th>Studio</th>
<th>ABC</th>
<th>NBC</th>
<th>FOX</th>
<th>CBS</th>
<th>All broadcast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walt Disney Co</td>
<td>13.3%</td>
<td>11.3%</td>
<td>3.9%</td>
<td>6.5%</td>
<td>35.0%</td>
</tr>
<tr>
<td>Comcast Corp</td>
<td>7.7%</td>
<td>9.9%</td>
<td>4.0%</td>
<td>6.8%</td>
<td>28.4%</td>
</tr>
<tr>
<td>News Corp</td>
<td>9.4%</td>
<td>6.6%</td>
<td>6.3%</td>
<td>6.0%</td>
<td>28.3%</td>
</tr>
<tr>
<td>Time Warner Inc</td>
<td>12.6%</td>
<td>11.6%</td>
<td>7.9%</td>
<td>12.0%</td>
<td>44.1%</td>
</tr>
<tr>
<td>Sony Corp</td>
<td>7.0%</td>
<td>8.5%</td>
<td>5.2%</td>
<td>5.6%</td>
<td>26.3%</td>
</tr>
</tbody>
</table>
different studios allocate towards broadcast advertising (like Warner vs Sony), which may lead to spurious effects. To overcome that problem, we use the allocation to other networks as a control and estimate a Difference-in-Differences specification, where $i$ again indexes the products and $j$ indexes the networks:

$$s_{ij} = \delta + \beta_1 AffiliatedNetwork_j + \beta_2 AffiliatedStudio_i + \beta_3 AffiliatedStudio_i \times AffiliatedNetwork_j + \varepsilon_{ij}$$  \hspace{1cm} (3.19)

We estimate two versions of (3.19): (i) taking all non-focal broadcasters as a control and (ii) taking only CBS, as the non-affiliated network, as a control. In this specification, we (i) take the difference between the share of ABC and CBS (or average across NBC, FOX and CBS) for the control products and the Disney movies; and (ii) compare the average difference Disney to the average difference for control movies to find the effect of interest, which is captured by $\beta_3$. In this case, we allow for some movies to have a higher overall level of advertising on broadcast, which is differenced out.

Connecting these regressions to our theory model, Table 12 documents the difference between $A^*_D$ and $A^*_TW$. However, it may stem from both $\lambda > 0$ and $\bar{u}_D > \bar{u}_TW$. We use matching to select products that have $\bar{u}_i \approx \bar{u}_D$. Then if for those products $A^*_i < A^*_D$, we can conclude that $\lambda > 0$.

3.4.1. Matching Methodology

For each movie produced by Comcast, Walt Disney Co or News Corp (i.e. including the major and the minor studios listed in Table 10), we select 20 products that are likely to have the most similar advertising strategy to the movie.

For the matching procedure, we use advertising data for cable networks that are not affiliated with either Comcast, Walt Disney or News Corp. There are 50 such networks in our data. They account for 46.6% of all TV advertising during the sample period. Table 13 lists the 20 most popular networks out of the set of 50. The table lists the share of all advertising and the share of all movie advertising within the set of unaffiliated networks. It also shows the
standard deviation of each network’s advertising share across all products and all movies. Movies do not necessarily follow other products’ advertising allocation. The variation in shares across movies is much higher than across other products. This can be explained by the fact that movies in particular are a very heterogeneous category and different movies can target very different audiences and therefore advertise on different channels.

We can see from descriptions provided that the set of channels is quite diverse, so we can expect the selection of where to allocate advertising across these channels to be informative. We also want to see if this diversity is reflected in advertising data. For each channel, we find the most characteristic advertiser category and film. We define this in the following way. First, we compute how much advertising each product category accounts for across all 50 selected networks. Second, we do the same within each network. Then for each network we find the category which has the highest difference between average share of advertising and their share of advertising within that network. The idea is that product categories, as well as movies, differ widely in how much they advertise overall, so if we just look at the top advertisers, for almost any channel they are going to be cars, insurance companies or telecoms; the top advertising movies are going to be blockbusters. It tells us more about the set of networks to see which products disproportionately advertise on each network. The results seem to match a priori intuition: toys advertise on Cartoon Network, Home Improvement on HGTV, pet food on Animal Planet etc.

The matching procedure goes as follows for a movie $m$:

1. Select the dates during which movie $m$ advertises

2. For each product advertising during this period (including $m$), compute the advertising allocation across the set of unaffiliated cable networks

3. Compute the quadratic distance between the vector of ad shares for $m$ with every other product
<table>
<thead>
<tr>
<th>Network</th>
<th>s all</th>
<th>s movies</th>
<th>sd s all</th>
<th>sd s movies</th>
<th>highest category</th>
<th>highest movie</th>
<th>Genre</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBS</td>
<td>6%</td>
<td>11%</td>
<td>2%</td>
<td>12%</td>
<td>Insurance</td>
<td>This Is 40</td>
<td>Comedy</td>
<td>Comedy, Sports</td>
</tr>
<tr>
<td>TOON</td>
<td>6%</td>
<td>9%</td>
<td>3%</td>
<td>11%</td>
<td>Toys</td>
<td>Walking With Dinosaurs</td>
<td>Animation/family</td>
<td>Cartoons</td>
</tr>
<tr>
<td>TNT</td>
<td>6%</td>
<td>10%</td>
<td>3%</td>
<td>10%</td>
<td>Insurance</td>
<td>Paranoia (2013)</td>
<td>Drama</td>
<td>Drama, Sports</td>
</tr>
<tr>
<td>HST</td>
<td>5%</td>
<td>4%</td>
<td>2%</td>
<td>8%</td>
<td>Insurance</td>
<td>Paranoia (2013)</td>
<td>Drama</td>
<td>History</td>
</tr>
<tr>
<td>HGTV</td>
<td>5%</td>
<td>2%</td>
<td>2%</td>
<td>6%</td>
<td>Home Improvement</td>
<td>Paranoia (2013)</td>
<td>Drama</td>
<td>Home and garden reality</td>
</tr>
<tr>
<td>TVLD</td>
<td>4%</td>
<td>2%</td>
<td>1%</td>
<td>5%</td>
<td>Insurance</td>
<td>Beautiful Creatures (2013)</td>
<td>Other</td>
<td>Classic TV</td>
</tr>
<tr>
<td>CNN</td>
<td>4%</td>
<td>3%</td>
<td>2%</td>
<td>8%</td>
<td>Finance</td>
<td>Olympus Has Fallen</td>
<td>Action</td>
<td>News</td>
</tr>
<tr>
<td>AE</td>
<td>4%</td>
<td>4%</td>
<td>1%</td>
<td>6%</td>
<td>Fast Food</td>
<td>This Is 40</td>
<td>Comedy</td>
<td>Reality, true crime</td>
</tr>
<tr>
<td>AMC</td>
<td>4%</td>
<td>6%</td>
<td>2%</td>
<td>12%</td>
<td>Pharma</td>
<td>Paranoia (2013)</td>
<td>Drama</td>
<td>Movies, prestige TV drama</td>
</tr>
<tr>
<td>FOOD</td>
<td>4%</td>
<td>2%</td>
<td>2%</td>
<td>6%</td>
<td>Prepared Dinners</td>
<td>Hotel Transylvania</td>
<td>Animation/family</td>
<td>Cooking reality</td>
</tr>
<tr>
<td>DSC</td>
<td>3%</td>
<td>4%</td>
<td>2%</td>
<td>5%</td>
<td>Cars</td>
<td>Rise Of The Planet Of The Apes</td>
<td>Action</td>
<td>Pop science documentaries</td>
</tr>
<tr>
<td>SPKE</td>
<td>2%</td>
<td>4%</td>
<td>1%</td>
<td>9%</td>
<td>Fast Food</td>
<td>Hansel Gretel Witch Hunters</td>
<td>Action</td>
<td>action, reality</td>
</tr>
<tr>
<td>APL</td>
<td>2%</td>
<td>1%</td>
<td>1%</td>
<td>7%</td>
<td>Pet Food</td>
<td>Walking With Dinosaurs</td>
<td>Family</td>
<td>animal documentaries</td>
</tr>
<tr>
<td>BET</td>
<td>2%</td>
<td>4%</td>
<td>1%</td>
<td>9%</td>
<td>Motion Pictures</td>
<td>Hansel Gretel Witch Hunters</td>
<td>Action</td>
<td>Comedy, News, Music</td>
</tr>
<tr>
<td>CMT</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>2%</td>
<td>Diapers</td>
<td>Beautiful Creatures (2013)</td>
<td>Other</td>
<td>Country music, reality</td>
</tr>
<tr>
<td>VHI</td>
<td>1%</td>
<td>2%</td>
<td>1%</td>
<td>5%</td>
<td>Fast Food</td>
<td>Hall Pass</td>
<td>Comedy</td>
<td>Music, reality</td>
</tr>
<tr>
<td>TEENNCK</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>5%</td>
<td>Video Games</td>
<td>Hotel Transylvania</td>
<td>Animation/family</td>
<td>Teen shows</td>
</tr>
<tr>
<td>TRAVEL</td>
<td>1%</td>
<td>6%</td>
<td>1%</td>
<td>1%</td>
<td>Travel</td>
<td>Hotel Transylvania</td>
<td>Animation/family</td>
<td>Reality travel</td>
</tr>
<tr>
<td>NFLNET</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>6%</td>
<td>Sport League</td>
<td>Avengers (2012)</td>
<td>Action</td>
<td>Sports</td>
</tr>
<tr>
<td>SCIENCE</td>
<td>1%</td>
<td>6%</td>
<td>1%</td>
<td>3%</td>
<td>Technology</td>
<td>John Carter</td>
<td>Action</td>
<td>Pop science documentaries</td>
</tr>
</tbody>
</table>

Table 13: Top 20 cable networks used for matching
We perform step (1) to control for programming availability and variation across networks over time.

To summarize the results of the matching procedure, we combine all matched product of movies by parent company (Table 14) and genre (Table 15) and see what share of those products belong to each of the product categories. First, note that there is a “long tail” of categories, which implies that there is little overlap in matched products and categories across movies even by the same studio or within the same genre. Second, the top 4 categories are always the same: Fast Food, Motion Pictures, Cars and Other. This is not necessarily driven by ubiquity of advertising by those industries. Table 14 also presents the top 20 categories by the total amount of TV advertising. It is remarkable how many heavily advertising categories (Insurance, Pharma, Home Improvement, Toiletries) are rarely matched to movies. Movies very often match with other movies, which suggests that there is something specific about movie advertising demand.6

6See Appendix for a version of this analysis restricted to only movies as potential control group.
<table>
<thead>
<tr>
<th>Category</th>
<th>Share</th>
<th>Category</th>
<th>Share</th>
<th>Category</th>
<th>Share</th>
<th>Category</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast Food</td>
<td>34%</td>
<td>Cars</td>
<td>14%</td>
<td>Motion Pictures</td>
<td>14%</td>
<td>Motion Pictures</td>
<td>14%</td>
</tr>
<tr>
<td>Motion Pictures</td>
<td>19%</td>
<td>Other</td>
<td>9%</td>
<td>Fast Food</td>
<td>13%</td>
<td>Fast Food</td>
<td>12%</td>
</tr>
<tr>
<td>Cars</td>
<td>7%</td>
<td>Other</td>
<td>9%</td>
<td>Other</td>
<td>10%</td>
<td>Other</td>
<td>7%</td>
</tr>
<tr>
<td>Other</td>
<td>4%</td>
<td>Cars</td>
<td>8%</td>
<td>Telecom</td>
<td>3%</td>
<td>Cars</td>
<td>7%</td>
</tr>
<tr>
<td>Telecom</td>
<td>3%</td>
<td>OTC meds</td>
<td>4%</td>
<td>Media</td>
<td>4%</td>
<td>Telecom</td>
<td>3%</td>
</tr>
<tr>
<td>Confectionery</td>
<td>3%</td>
<td>Telecom</td>
<td>3%</td>
<td>Confectionery</td>
<td>3%</td>
<td>Motion Pictures</td>
<td>3%</td>
</tr>
<tr>
<td>Media</td>
<td>3%</td>
<td>Motion Pictures</td>
<td>19%</td>
<td>Cleaning Products</td>
<td>3%</td>
<td>Cars</td>
<td>7%</td>
</tr>
<tr>
<td>Snacks</td>
<td>3%</td>
<td>Credit Cards</td>
<td>2%</td>
<td>Other</td>
<td>9%</td>
<td>Media</td>
<td>3%</td>
</tr>
<tr>
<td>Tablets \ Phones</td>
<td>2%</td>
<td>Prepared Dinners</td>
<td>2%</td>
<td>Other</td>
<td>9%</td>
<td>Prepared Dinners</td>
<td>2%</td>
</tr>
<tr>
<td>Razors</td>
<td>2%</td>
<td>Finance</td>
<td>2%</td>
<td>Confectionery</td>
<td>2%</td>
<td>Apparel</td>
<td>2%</td>
</tr>
<tr>
<td>Gum \ Mints</td>
<td>1%</td>
<td>Telecom</td>
<td>2%</td>
<td>Confectionery</td>
<td>2%</td>
<td>Tablets \ Phones</td>
<td>2%</td>
</tr>
<tr>
<td>Hair Products</td>
<td>1%</td>
<td>Tablets \ Phones</td>
<td>2%</td>
<td>Snacks</td>
<td>3%</td>
<td>Tablets \ Phones</td>
<td>2%</td>
</tr>
<tr>
<td>Snacks</td>
<td>1%</td>
<td>Travel</td>
<td>2%</td>
<td>Cream</td>
<td>2%</td>
<td>Snacks</td>
<td>3%</td>
</tr>
<tr>
<td>Insurance</td>
<td>1%</td>
<td>Toiletries</td>
<td>2%</td>
<td>Finance</td>
<td>2%</td>
<td>Cream</td>
<td>2%</td>
</tr>
<tr>
<td>Toiletries</td>
<td>1%</td>
<td>Gum \ Mints</td>
<td>2%</td>
<td>Confectionery</td>
<td>1%</td>
<td>Insurance</td>
<td>2%</td>
</tr>
<tr>
<td>Casual Dining</td>
<td>1%</td>
<td>Confectionery</td>
<td>1%</td>
<td>Finance</td>
<td>2%</td>
<td>Reckless</td>
<td>2%</td>
</tr>
<tr>
<td>Video Games</td>
<td>1%</td>
<td>Home Improvement</td>
<td>1%</td>
<td>Confectionery</td>
<td>1%</td>
<td>Reckless</td>
<td>2%</td>
</tr>
<tr>
<td>OTC meds</td>
<td>1%</td>
<td>Weight Loss</td>
<td>1%</td>
<td>OTC meds</td>
<td>3%</td>
<td>Reckless</td>
<td>2%</td>
</tr>
<tr>
<td>Deodorant men</td>
<td>1%</td>
<td>Weight Loss</td>
<td>1%</td>
<td>OTC meds</td>
<td>3%</td>
<td>Reckless</td>
<td>2%</td>
</tr>
<tr>
<td>Tech Stores</td>
<td>1%</td>
<td>Insurance</td>
<td>1%</td>
<td>OTC meds</td>
<td>3%</td>
<td>Reckless</td>
<td>2%</td>
</tr>
<tr>
<td>Computers</td>
<td>1%</td>
<td>Skin Care</td>
<td>1%</td>
<td>OTC meds</td>
<td>3%</td>
<td>Reckless</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Snacks</td>
<td>1%</td>
<td>OTC meds</td>
<td>3%</td>
<td>Reckless</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 15: Top 20 categories of matched products by movie genre

3.4.2. Results

We take the matched products identified in the previous subsection and use them to run regressions (3.18) and (3.19). Tables 16, 17 and 18 present the results for NBC-Universal, ABC-Disney and FOX-20th Century respectively.

In the first three columns, the sample consists of all movies released by studios owned by the conglomerate, both major and minor. In columns OLS2 to DID4, we use only major studio films, i.e. only movies by Universal Pictures, Walt Disney Pictures and 20th Century Fox Pictures. In columns OLS1 and OLS2, we estimate a model described by equation (3.18), which ignores the differences in the overall level of broadcast advertising across studios. In columns DID1 and DID3, we estimate equation (3.19) using shares of all 4 major broadcast networks as DV. In other words, to estimate the affiliation effect for NBC-Universal, we use ABC, FOX and CBS as control group for NBC. Note that since ABC and FOX are affiliated with competing movie studios, they have an incentive to charge Universal high prices for advertising, which may bias the estimates of the affiliation effect in positive direction. To
Table 16: Regression results for Comcast movies and matched products.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBC</td>
<td>0.021*** (0.002)</td>
<td>0.009*** (0.003)</td>
<td>0.020*** (0.003)</td>
<td>0.010*** (0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal</td>
<td>0.005 (0.005)</td>
<td>-0.006 (0.010)</td>
<td>0.008 (0.006)</td>
<td>-0.004 (0.011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBCXUniversal</td>
<td>0.007 (0.011)</td>
<td>0.001 (0.010)</td>
<td>0.013 (0.014)</td>
<td>0.008 (0.012)</td>
<td>0.000 (0.012)</td>
<td>0.012 (0.016)</td>
</tr>
<tr>
<td>Nobs</td>
<td>1617</td>
<td>6468</td>
<td>3234</td>
<td>1176</td>
<td>4704</td>
<td>2352</td>
</tr>
<tr>
<td>Nparams</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.000</td>
<td>0.014</td>
<td>0.003</td>
<td>0.000</td>
<td>0.013</td>
<td>0.003</td>
</tr>
</tbody>
</table>

overcome this problem, we use only the unaffiliated CBS as a control group in regressions DID2 and DID4. The affiliation effect (i.e. the effect of interest) is captured by the studio-network interaction.

We find no significant affiliation effect for NBC-Universal. This is the most recent relationship out of the ones we study, established only 7 years before the sample period. During the sample period, NBC-Universal is going through a merger with Comcast, which could also negatively affect the coordination between different divisions.

We find significant affiliation effects for ABC-Disney and FOX-20th Century. The effects are directionally stronger for the major studio releases, however the difference is not big enough to be statistically significant. We find that due to affiliation, Disney movies allocate between 4.5 and 8.9 percentage points more advertising on ABC and 20th Century movies allocate between 2 and 6.8 percentage points more towards FOX.

3.4.3. Falsification Exercise

We want to establish that or statistical test of an integrated relationship does not create a spurious effect. To do that, we perform the same matching exercise and run the same regressions as in the previous section under the false assumption of affiliation between unaffiliated studios (Warner Bros, Paramount and Sony) and the three networks we are interested in. The results are reported in Tables 19, 20 and 21. In case of Warner, we find
Table 17: Regression results for Disney movies and matched products.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>0.012***</td>
<td>-0.005</td>
<td>0.010***</td>
<td>-0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disney</td>
<td>0.015**</td>
<td>-0.026**</td>
<td>0.010</td>
<td>-0.025**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.012)</td>
<td>(0.007)</td>
<td>(0.012)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCXDisney</td>
<td>0.060***</td>
<td>0.045***</td>
<td>0.086***</td>
<td>0.064***</td>
<td>0.054***</td>
<td>0.089***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.017)</td>
<td>(0.012)</td>
<td>(0.014)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>Nobs</td>
<td>924</td>
<td>3696</td>
<td>1848</td>
<td>798</td>
<td>3192</td>
<td>1596</td>
</tr>
<tr>
<td>Nparams</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.026</td>
<td>0.013</td>
<td>0.014</td>
<td>0.033</td>
<td>0.013</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Table 18: Regression results for News Corp movies and matched products.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOX</td>
<td>-0.040***</td>
<td>-0.036***</td>
<td>-0.036***</td>
<td>-0.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20thCentury</td>
<td>0.012**</td>
<td>-0.003</td>
<td>-0.011*</td>
<td>-0.025**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.009)</td>
<td>(0.006)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOXX20thCentury</td>
<td>0.032***</td>
<td>0.020*</td>
<td>0.035***</td>
<td>0.043***</td>
<td>0.053***</td>
<td>0.068***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.008)</td>
<td>(0.012)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Nobs</td>
<td>1533</td>
<td>6132</td>
<td>3066</td>
<td>1071</td>
<td>4284</td>
<td>2142</td>
</tr>
<tr>
<td>Nparams</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.014</td>
<td>0.044</td>
<td>0.055</td>
<td>0.024</td>
<td>0.038</td>
<td>0.047</td>
</tr>
</tbody>
</table>
significant effects only when we estimate specification (3.18), due to Warner’s strategy to allocate more of its advertising towards broadcast TV than other studios. When we control for that, the effect disappears. For Sony (Columbia), none of the “affiliation” effects are robustly significant. Finally, one can see in Table 20 that we find a significant relationship between FOX and Paramount, despite them not being affiliated in reality. This is likely to happen by chance due to the fact that we are testing multiple hypotheses.

Table 19: Falsification exercise: estimated effects from treating Warner Bros as an affiliated studio.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBCXWarner</td>
<td>0.017*</td>
<td>-0.019*</td>
<td>-0.007</td>
<td>0.022**</td>
<td>-0.017*</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.013)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>ABCXWarner</td>
<td>0.043***</td>
<td>0.015</td>
<td>0.018</td>
<td>0.045***</td>
<td>0.014</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.010)</td>
<td>(0.012)</td>
<td>(0.008)</td>
<td>(0.010)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>FOXXWarner</td>
<td>0.042***</td>
<td>0.014</td>
<td>0.017</td>
<td>0.044***</td>
<td>0.013</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.006)</td>
<td>(0.010)</td>
<td>(0.011)</td>
</tr>
</tbody>
</table>

Table 20: Falsification exercise: estimated effects from treating Paramount as an affiliated studio.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBCXParamount</td>
<td>-0.008</td>
<td>-0.016</td>
<td>0.010</td>
<td>-0.011</td>
<td>-0.019</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.015)</td>
<td>(0.011)</td>
<td>(0.013)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>ABCXParamount</td>
<td>0.004</td>
<td>0.000</td>
<td>0.022</td>
<td>0.003</td>
<td>-0.001</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.012)</td>
<td>(0.015)</td>
<td>(0.010)</td>
<td>(0.013)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>FOXXParamount</td>
<td>0.037***</td>
<td>0.045***</td>
<td>0.055***</td>
<td>0.040***</td>
<td>0.048***</td>
<td>0.057***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.012)</td>
<td>(0.015)</td>
<td>(0.010)</td>
<td>(0.012)</td>
<td>(0.015)</td>
</tr>
</tbody>
</table>

Table 21: Falsification exercise: estimated effects from treating Columbia as an affiliated studio.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBCXColumbia</td>
<td>-0.002</td>
<td>-0.015</td>
<td>-0.022</td>
<td>-0.011</td>
<td>-0.009</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.015)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>ABCXColumbia</td>
<td>0.010</td>
<td>0.001</td>
<td>-0.011</td>
<td>0.005</td>
<td>0.012</td>
<td>0.028**</td>
</tr>
<tr>
<td></td>
<td>(0.010)</td>
<td>(0.011)</td>
<td>(0.015)</td>
<td>(0.010)</td>
<td>(0.012)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>FOXXColumbia</td>
<td>0.008</td>
<td>-0.001</td>
<td>-0.012</td>
<td>0.012</td>
<td>0.021*</td>
<td>0.035***</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.012)</td>
<td>(0.008)</td>
<td>(0.012)</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>
Table 22: Regression results for Comcast movies and matched products, heterogeneous effect for 10% most heavily advertising films.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBCXUniversal</td>
<td>0.000</td>
<td>-0.005</td>
<td>0.007</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.015)</td>
</tr>
<tr>
<td>NBCXUniversal:blockbuster</td>
<td>0.047</td>
<td>0.045</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.031)</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Nobs</td>
<td>1638</td>
<td>6552</td>
<td>3276</td>
</tr>
<tr>
<td>Nparams</td>
<td>2</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.000</td>
<td>0.016</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Table 23: Regression results for Walt Disney movies and matched products, heterogeneous effect for 10% most heavily advertising films.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCXDisney</td>
<td>0.062***</td>
<td>0.044***</td>
<td>0.086***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.015)</td>
<td>(0.018)</td>
</tr>
<tr>
<td>ABCXDisney:blockbuster</td>
<td>-0.018</td>
<td>0.007</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.044)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>Nobs</td>
<td>924</td>
<td>3696</td>
<td>1848</td>
</tr>
<tr>
<td>Nparams</td>
<td>3</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.024</td>
<td>0.015</td>
<td>0.014</td>
</tr>
</tbody>
</table>

3.4.4. Treatment Effect Heterogeneity

Another way to test our approach is to consider heterogeneity in treatment effects. If the effect that we estimate captures only the affiliation between the studio and the network ($\lambda$), then there cannot be any significant heterogeneity in treatment effects across movies produced by the same studio. We consider two dimensions of movie heterogeneity. First, we separate the top 10% most highly advertised movies and see if their relationship with the affiliated network differs from the other 90% of the movies. Tables 22, 23 and 24 show the estimation results. We find no significant heterogeneity across the board.

Second, we consider genre heterogeneity. Can the affiliation effect be stronger for Disney animation movies compared to all other Disney movies? For considerations of statistical power, we limit ourselves to a maximum of two top genres per studio. Those genres are: comedy and action for Comcast, animation for Disney, comedy and drama for News Corp.
Table 24: Regression results for News Corp movies and matched products, heterogeneous effect for 10% most heavily advertising films.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOXX20thCentury</td>
<td>0.029</td>
<td>0.015</td>
<td>0.033</td>
<td>0.006</td>
<td>0.011</td>
<td>0.012</td>
</tr>
<tr>
<td>FOXX20thCentury:blockbuster</td>
<td>-0.028</td>
<td>0.023</td>
<td>0.000</td>
<td>0.023</td>
<td>0.038</td>
<td>0.043</td>
</tr>
<tr>
<td>Nobs</td>
<td>1617</td>
<td>6468</td>
<td>3234</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nparams</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R2adj</td>
<td>0.012</td>
<td>0.045</td>
<td>0.055</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 25: Regression results for Comcast films, with treatment heterogeneity for top two movie genres produced by the studio: comedy and action.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBCXUniversal</td>
<td>-0.002</td>
<td>-0.002</td>
<td>0.017</td>
<td>0.004</td>
<td>0.000</td>
<td>0.020</td>
</tr>
<tr>
<td>NBCXUniversal:comedy</td>
<td>0.008</td>
<td>0.007</td>
<td>-0.008</td>
<td>0.002</td>
<td>0.004</td>
<td>-0.012</td>
</tr>
<tr>
<td>NBCXUniversal:action</td>
<td>0.016</td>
<td>0.001</td>
<td>-0.005</td>
<td>0.009</td>
<td>-0.005</td>
<td>-0.010</td>
</tr>
<tr>
<td>Nobs</td>
<td>1638</td>
<td>6552</td>
<td>3276</td>
<td>1176</td>
<td>4704</td>
<td>2352</td>
</tr>
<tr>
<td>Nparams</td>
<td>3</td>
<td>11</td>
<td>11</td>
<td>3</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>R2adj</td>
<td>-0.001</td>
<td>0.013</td>
<td>0.002</td>
<td>-0.002</td>
<td>0.013</td>
<td>0.000</td>
</tr>
</tbody>
</table>

We find no significant heterogeneity for Comcast (Table 25) and Disney (Table 26) movies. Surprisingly, we find significant heterogeneity for comedy and drama produced by News Corp-affiliated studios in the first three columns of Table 27. However, if we focus only on movies produced by the major studio, 20th Century Fox, the genre heterogeneity is not significant. We interpret these results in the following way. 20th Century Fox and Fox Searchlight differ in their level of integration with FOX, specifically Fox Searchlight is more independent and less coordinated with the rest of the conglomerate. The studios differ in the genre specialization. Fox Searchlight produces mostly drama and comedy movies, so the differences between the minor and the major studios are reflected through genre heterogeneity.
### Table 26: Regression results for Disney films, with treatment heterogeneity for animation and family movies.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABCXDisney</td>
<td>0.077***</td>
<td>0.058***</td>
<td>0.089***</td>
<td>0.078***</td>
<td>0.075***</td>
<td>0.090***</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.020)</td>
<td>(0.025)</td>
<td>(0.017)</td>
<td>(0.020)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>ABCXDisney:animation</td>
<td>-0.034</td>
<td>-0.026</td>
<td>-0.007</td>
<td>-0.027</td>
<td>-0.043</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.028)</td>
<td>(0.034)</td>
<td>(0.024)</td>
<td>(0.029)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Nobs</td>
<td>924</td>
<td>3696</td>
<td>1848</td>
<td>798</td>
<td>3192</td>
<td>1596</td>
</tr>
<tr>
<td>Nparams</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.033</td>
<td>0.015</td>
<td>0.019</td>
<td>0.036</td>
<td>0.013</td>
<td>0.018</td>
</tr>
</tbody>
</table>

### Table 27: Regression results for News Corp films, with treatment heterogeneity for comedy and drama movies - top two genres produced by the studio.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOXX20thCentury</td>
<td>0.046***</td>
<td>0.062***</td>
<td>0.076***</td>
<td>0.050***</td>
<td>0.066***</td>
<td>0.078***</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.016)</td>
<td>(0.018)</td>
<td>(0.010)</td>
<td>(0.015)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>FOXX20thCentury:comedy</td>
<td>-0.037***</td>
<td>-0.095***</td>
<td>-0.087***</td>
<td>-0.036*</td>
<td>-0.049</td>
<td>-0.046</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.026)</td>
<td>(0.030)</td>
<td>(0.020)</td>
<td>(0.031)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>FOXX20thCentury:drama</td>
<td>-0.024</td>
<td>-0.072***</td>
<td>-0.071**</td>
<td>0.009</td>
<td>-0.022</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.026)</td>
<td>(0.030)</td>
<td>(0.026)</td>
<td>(0.040)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Nobs</td>
<td>1554</td>
<td>6216</td>
<td>3108</td>
<td>1071</td>
<td>4284</td>
<td>2142</td>
</tr>
<tr>
<td>Nparams</td>
<td>5</td>
<td>11</td>
<td>11</td>
<td>5</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.018</td>
<td>0.050</td>
<td>0.059</td>
<td>0.039</td>
<td>0.040</td>
<td>0.051</td>
</tr>
</tbody>
</table>
3.5. Conclusion

We analyze theoretically and empirically the implications of integration between major film studios and broadcast television networks. We construct a theoretical model of a duopolistic movie market, in which studios compete in persuasive advertising levels, which are mediated through a monopolistic TV network. One of the studios is integrated with the TV network. We find that increasing integration leads to higher and more optimal level of advertising and higher market share for the integrated movie studio and lower advertising levels and market share for the non-integrated studio. From the perspective of total welfare, we find that there is an optimal level of integration that is between 0 (the studio is independent from the network) and 1 (the studio and network are fully integrated).

Our theoretical model also shows us that there are two forces that can lead a studio to disproportionately advertise on a given network: (i) content match between the studio’s films and the network’s programming (i.e. the movies and the shows are serving the same target audience) and (ii) vertical integration. Our empirical exercise is focused on separating the two forces and documenting the effect of affiliation on advertising demand. To achieve this goal, we combine a Difference-in-Differences empirical structure with a matching procedure to make sure that we select the control group of products that have as close as possible content match with the treated movie.

We use data from Rentrak Corporation to estimate the affiliation effects. We find significant affiliation effects for the relationships between ABC and Disney (between 4.5 and 8.9 percentage points increase in share of ABC) and FOX and News Corp (between 2 and 6.8 percentage points increase in share of FOX). We find no significant effect for NBC-Universal, however, this relationship is only about 7 years old and is going through another corporate transformation (acquired by Comcast) during the sample period, so it is not unreasonable to believe that the desired level of integration for the conglomerate has not been achieved yet. If we exclude movies by minor studios affiliated with Disney, 20th Century and Universal from the analysis, the affiliation effects get stronger. It is particularly pronounced for
20th Century Fox and Fox Searchlight, which points to the idea that potentially the major studio is more integrated with the rest of the conglomerate than the minor.

We perform a series of robustness checks. First, we perform a falsification exercise in which we follow the same procedure as in our main estimation under the false assumption that Warner Bros, Paramount or Columbia Pictures are integrated with one of either ABC, NBC or FOX. Our procedure returns one false positive out of the nine false relationships we test, making it unlikely that both our findings for ABC-Disney and FOX-News Corp are false positives. Second, we try to estimate treatment effect heterogeneity across movie advertising budgets and genres. If we do not control well for content match with our matching procedure, it is likely there would be significant heterogeneity in affiliation effects, as different films would have different content match with the affiliated network. We find no evidence of significant heterogeneity.

The broader implications of our findings go beyond the context of movie or TV advertising market. The main takeaway is that vertical integration between a platform and a market participant in a two-sided market can lead to the integrated market seller’s products to be featured more prominently on the platform, independent of whether that product is the best match for consumers. This has further implications for competition and welfare. From competition perspective, the competitor of the vertically integrated firm loses market share. From welfare perspective, even if the product by the integrated firm is better in a vertical sense, integration can reduce welfare by squeezing the competitor out of the market.
APPENDIX for Chapter 1

A.1. Proofs

Proof of Proposition 1.

Because the profits of sellers 1 and 2 are ex-ante symmetric from the viewpoint of the platform, the optimal prices will have \( p^*_1 = p^*_2 \). The solutions to the first order conditions on price yield the expressions in the proposition. When setting \( p = p_1 = p_2 \), the expected profit is concave in price, hence the solution is a unique equilibrium.

The comparative statics analysis of prices follow from the linearity of optimal prices in all parameters.

For \( \alpha \), \( \frac{\partial \pi^{CP}}{\partial \alpha} = -2(v + 1 - 2\alpha) < 0 \) because \( \alpha < 1/2 \). Integrating over \( v \) and using Leibniz’s integral rule also proves that \( \frac{\partial \mathbb{E}(\pi^{CP})}{\partial \alpha} < 0 \). \( \square \)

Proof of Proposition 2.

To find the equilibrium prices, the FOC of a seller of type \( \tau \) is:

\[
\frac{2\mathbb{E}(v) + q(3 - 2\alpha) + 2 - 4\alpha}{2} - p^C_r(2\alpha - 3) + (p^C_L + p^C_H) \frac{1 - 2\alpha}{4} = 0 \quad (A.1)
\]

Imposing \( p^*_r = p^C_r \) results in the equilibrium prices as the solution. Comparative statics with respect to \( q \) follow from the linearity in \( q \).

For \( \alpha \), \( \frac{\partial p^C_r}{\alpha} = \frac{4(\mathbb{E}(v)-4)}{(5-2\alpha)^2} > 0 \) when \( \mathbb{E}(v) > 4 \). \( \square \)

The other items result from plugging-in the prices into the profit functions and integrating over values where necessary.

Proof of Corollary 1.
Comparing $p^C_L$ to $p^*$, we find that $p^C_L > p^*$ when

$$v < -q + \frac{2(\pi + v) - (1 - 2\alpha)^2}{5 - 2\alpha}$$

and

$$\bar{v} > \frac{1}{2} \left( q(5 - 2\alpha) + (1 - 2\alpha)^2 + (3 - 2\alpha)v \right)$$

For the second item, the platform’s profit $\pi^C_C$ is compared to $\pi^C_P$. Because $\pi^C_P$ is quadratic and increasing in $q$, and because $\pi^C_P$ does not depend on $q$, finding the $q$ for which $\pi^C_C = \pi^C_P$ gives the solution in the proposition. Finally, because $\pi^C_C \bigg|_{q=0} \leq \pi^C_P$, there is a crossover of profits as described in the proposition.

For the third item, we follow a similar approach to the second item. $E(\pi^C_C)$ is increasing and quadratic in $q$ and $E(\pi^C_P)$ is a constant, therefore, to show the existence of a crossing point, we only need to show that there is a point such that $E(\pi^C_C) < E(\pi^C_P)$ for some $q$. We take the lowest possible value, $q = 0$. Then the inequality can be reduced to the following:

$$\frac{4(3 - 2\alpha)(E(v) + 1 - 2\alpha)^2}{(5 - 2\alpha)^2} < E\left(\frac{(v + 1 - 2\alpha)^2}{2}\right)$$

Note that (i) $\frac{4(3 - 2\alpha)}{(5 - 2\alpha)^2} < \frac{1}{2}$ and (ii) $(E(v) + 1 - 2\alpha)^2 < E((v + 1 - 2\alpha)^2)$ by Jensen’s inequality.

For the fourth term, first observe that $E(\pi^C_H)$ is a linear increasing function of $q$ and $E(\pi^C_C)$ is a quadratic increasing function of $q$. Moreover, if $q = 0$, $E(\pi^C_H) = (1 - \phi) \frac{E(\pi^C_C)}{2\phi} > \frac{E(\pi^C_C)}{2\phi} = E(\pi^C_C)$. Second, we can rewrite $E(\pi^C_C) = (1 - \phi)(\frac{E(\pi^C_C)}{2\phi} - \frac{(v + v + 2(1 - 2\alpha)q)}{2(5 - 2\alpha)})$. As $\frac{3 - 2\alpha}{4} > \frac{1}{5 - 2\alpha}$, $E(\pi^C_C)$ is decreasing in $q$ faster than $E(\pi^C_C)$. At $q = 0$ the profits are again proportional to the platform’s expected profit and therefore there is a crossing point.

**Proof of Lemma 1.**

Using the expressions for for prices $p^C_L$ and $p^C_H$, they depend on $v$ solely through the belief sellers have about $E(v)$. Hence, given any equilibrium strategy $m(v)$ and corresponding $m$,
the unique equilibrium prices are as specified in the text.

Proof of Lemma 2.

We follow the steps from Lemma 1 in Crawford and Sobel (1982). Assume w.l.o.g. that $E(v|m_1) < E(v|m_2)$. First, there exists a state $\tilde{v}$, such that $\pi^P(\tilde{v}, m_1) = \pi^P(\tilde{v}, m_2)$. This state is

$$\tilde{v} = \frac{2(E(v|m_1) + E(v|m_2)) - (1 - 2\alpha)^2}{5 - 2\alpha}$$  \hspace{1cm} (A.2)

and the optimal induced expectation in that state for the platform is $E^*(\tilde{v}) = \frac{E(v|m_1) + E(v|m_2)}{2} \in (E(v|m_1), E(v|m_2))$.

Second, it follows that $E(v|m_1)$ is not induced in equilibrium in any state $v > \tilde{v}$, as it is more profitable to induce $E(v|m_2)$ and vice versa, $E(v|m_2)$ is not induced in any state $v < \tilde{v}$.

Third, since $E(v|m_1)$ is a rational expectation over which states the platform would choose to induce such expectation, $E(v|m_1) \leq \tilde{v}$ and similarly $E(v|m_2) \geq \tilde{v}$.

Fourth, we know that $E^*(v) \neq v \forall v$, then $|\tilde{v} - \frac{E(v|m_1) + E(v|m_2)}{2}| > \epsilon$, which means that $E(v|m_2) - E(v|m_1) > \epsilon$. Since the state space is bounded, this means that there can only be a finite number of induced expectations.

Proof of Proposition 3.

First, we prove that price recommendation always gives a higher payoff than competition to both the platform and the sellers. Given that $\sum_{k=1}^{n} \frac{v_k - v_{k-1}}{\bar{v} - \bar{v}} \frac{v_k + v_{k-1}}{2} = E(v)$, the condition $E(\pi_t^R) > E(\pi_t^C)$ for $t \in \{P, H, L\}$ can be reduced to

$$\sum_{k=1}^{n} \frac{v_k - v_{k-1}}{\bar{v} - \bar{v}} \left( \frac{v_k + v_{k-1}}{2} \right)^2 > \left( \frac{\bar{v} + \bar{v}}{2} \right)^2$$

which is true by convexity of the function $f(x) = x^2$.

105
Second, we prove that when \( q = 0 \), \( \mathbb{E}(\pi^R_H) < \mathbb{E}(\pi^{CP}_H) \). This inequality reduces to

\[
\sum_{k=1}^{n} \frac{v_k - v_{k-1}}{\bar{v} - \underline{v}} \frac{4(3 - 2\alpha)((v_k + v_{k-1})/2 + t(1 - 2\alpha))^2}{(5 - 2\alpha)^2} < \frac{\mathbb{E}(v + t(1 - 2\alpha))^2}{2} \iff \sum_{k=1}^{n} \frac{v_k - v_{k-1}}{\bar{v} - \underline{v}} \frac{4(3 - 2\alpha)(\mathbb{E}(v|k) + t(1 - 2\alpha))^2}{(5 - 2\alpha)^2} \ll
\]

Note that (i) \( \frac{4(3 - 2\alpha)}{(5 - 2\alpha)^2} < \frac{1}{2} \) and (ii) \( (\mathbb{E}(v|k) + t(1 - 2\alpha))^2 < \mathbb{E}((v + t(1 - 2\alpha))^2|k) \) by Jensen’s inequality. Then the inequality holds. The rest of the proof follows the same steps as the proof of Corollary 1.

\[\square\]

**Proof of Corollary 2.**

We first prove that the cheap talk equilibria satisfy the monotonicity criterion that for every two cheap talk equilibria characterized by vectors of interval boundaries \( \hat{v} = (\hat{v}_0, \hat{v}_1, \ldots) \) and \( \tilde{v} = (\tilde{v}_0, \tilde{v}_1, \ldots) \) with \( \hat{v}_0 = \tilde{v}_0 = \underline{v} \) and \( \hat{v}_1 > \tilde{v}_1 \) then \( \hat{v}_i > \tilde{v}_i \) for every \( i \geq 2 \).

Let \( \hat{v} \) and \( \tilde{v} \) be solutions for the interval boundaries in Equation (1.15). By definition, \( \hat{v}_0 = \tilde{v}_0 = \underline{v} \). Suppose \( \hat{v}_1 > \tilde{v}_1 \) then \( \hat{v}_1 - \tilde{v}_1 = \frac{\hat{v}_2 - \tilde{v}_2}{3 - 2\alpha} > 0 \) which results in \( \hat{v}_2 > \tilde{v}_2 \). Now assume that monotonicity applies for every \( i < m \). Then: \( \hat{v}_{m-1} - \tilde{v}_{m-1} = \frac{\hat{v}_m - \tilde{v}_m}{3 - 2\alpha} > 0 \) and by the same logic from above, this proves that monotonicity holds by induction.

Using Proposition 3 from Chen et al. (2008), because the equilibria in the game are monotonic, only the unique equilibrium partition with the maximum number of induced actions satisfies NITS.

\[\square\]

**Proof of Proposition 4.**

First consider the high type. Note that \( \mathbb{E}(\pi^R_H) = (1 - \phi)\frac{\mathbb{E}(\pi^R_H)}{2\phi} + \frac{(\mathbb{E}(\pi^{CP}_H) + 1 - 2\alpha)q}{5 - 2\alpha} \). If \( q = \hat{q} \), \( \mathbb{E}(\pi^R_H) < \mathbb{E}(\pi^{CP}_H) \) as the first summands are equal, but the second linear part is larger for \( \mathbb{E}(\pi^{CP}_H) \). As the profit in case of recommendation is quadratic and increasing, the crossing point has to be further to the right than \( \hat{q} \). Then \( \hat{q}_H > \hat{q} \).
Second, consider the low type. Again we can write it as a combination of quadratic increasing function of $q$ and a linear decreasing function of $q$. \[ E(\pi^R_L) = (1 - \phi) \left( \frac{E(\pi^R_L)}{2\phi} - \frac{(E(\pi) + 1 + 2\alpha q)}{5 - 2\alpha} \right). \]

By the same logic as before, we get that at $q = \tilde{q}$, $E(\pi^R_L) > E(\pi^C_P)$. We also know that at $q = 0$ $E(\pi^R_L) < E(\pi^C_P)$. Both functions are monotonic, so it has to be the case that $\tilde{q}_L < \tilde{q}$.

The proof for the competition regime follows the same steps. \qed

**Proof of Proposition 5.**

As established in the text, the relationship between the expected total demands is determined by the expected prices. They are the following:

\[
\begin{align*}
E(p^{CP}(v)) &= \frac{\bar{v} + v + 2(1 - 2\alpha)}{4} \\
E(p^C) &= \frac{p^C_H + p^C_L}{2} = \frac{\bar{v} + v + 2(1 - 2\alpha)}{5 - 2\alpha} \\
E(p^R) &= \sum_{k=1}^{n} \frac{v_k - v_{k-1} + v_k - v_{k-1} + 2(1 - 2\alpha)}{\bar{v} - v} = \frac{\bar{v} + v + 2(1 - 2\alpha)}{5 - 2\alpha}
\end{align*}
\]

Higher expected prices result in lower expected demand. Because $5 - 2\alpha > 4$, the expected demand is ordered as specified in the text. \qed

**Proof of Proposition 6.**

First we need to derive the expressions for expected consumer surplus. Captives purchase from the seller that they are aware of if and only if their distance from that seller is below $v - p_i + q_i$. The maximum utility a captive buyer can achieve is $v - p_i + q_i$. The utility of captive buyers decreases linearly with distance. Then the total surplus of captive consumers is an area under a triangle with base $2(v - p_i + q_i)$ and height $v - p_i + q_i$. Then $CS^{cap} = (v - p_i + q_i)^2$.

For shoppers, those located at $(-M, -1) \cup (1, M)$ act as captives (by Assumption A1). Then their surplus is $\frac{(v-p_i+q_i)^2}{2}$. The shoppers in the interval $[-1, 1]$ choose between the
two sellers. Denote by \( \tilde{x} = \frac{q_1 - q_2 - p_1 + p_2}{2} \) the location of the shopper indifferent between the two sellers. Then the consumer surplus of the buyers in \([-1, \tilde{x}]\) is a trapezoid with area 
\[
(\tilde{x} + 1) \frac{2(v + q_1 - p_1) + (\tilde{x} + 1)}{2}
\]
and for the buyers in \([\tilde{x}, 1]\) it is 
\[
(1 - \tilde{x}) \frac{2(v + q_2 - p_2) + (1 - \tilde{x})}{2}.
\]
Combining everything together, we get the following expression for consumer surplus:

\[
\text{CS}(p_1, p_2) = (v + q_1 - p_1)^2 + (v + q_2 - p_2)^2 +
+(1 - 2\alpha) \left( \frac{2(v + q_1 - p_1)(\tilde{x} + 1) + 2(\tilde{x}^2 + 1) + 2(v + q_2 - p_2)(1 - \tilde{x})}{2} \right)
\]

Using equilibrium prices, the expected consumer surplus under centralized pricing is:

\[
\mathbb{E}(\text{CS}^{CP}) = \mathbb{E}(v^2) + q^2(5 - 2\alpha) - 3 + 8\alpha - 4\alpha^2
\]  
(A.3)

The expected consumer surplus under competition is

\[
\mathbb{E}(\text{CS}^C) = \frac{q^2(5 - 2\alpha)}{8} + \frac{1}{(5 - 2\alpha)^2} \left( 2\mathbb{E}(2\mathbb{E}(v) - v(5 - 2\alpha))^2 \right) + (1 - 2\alpha)^2(3 - 2\alpha)\mathbb{E}(v)
\]

\[
-37 + 126\alpha - 124\alpha^2 + 40\alpha^3
\]

And the expected consumer surplus under recommendation regime is

\[
\mathbb{E}(\text{CS}^R) = \frac{q^2(5 - 2\alpha)}{8} + \frac{1}{(5 - 2\alpha)^2} \left( 2\sum_{k=1}^{n} \frac{v_k - v_{k-1}}{\bar{v} - v} \mathbb{E}((2\mathbb{E}(v|k) - v(5 - 2\alpha))^2|k) +
\right.
\]

\[
+2(1 - 2\alpha)^2(3 - 2\alpha)\mathbb{E}(v) - 37 + 126\alpha - 124\alpha^2 + 40\alpha^3
\]
Second, we need to show that $\mathbb{E}(CS^C) > \mathbb{E}(CS^R)$. This reduces to

$$
\mathbb{E}((2\mathbb{E}(v) - v(5 - 2\alpha))^2) > \sum_{k=1}^{n} \frac{v_k - v_{k-1}}{v - \bar{v}} \mathbb{E}((2\mathbb{E}(v|k) - v(5 - 2\alpha))^2|k) \iff
$$

$$
-8(2 - \alpha)(\mathbb{E}(v))^2 > -8(2 - \alpha) \sum_{k=1}^{n} \frac{v_k - v_{k-1}}{v - \bar{v}} (\mathbb{E}(v|k))^2 \iff
$$

$$
-8(2 - \alpha) \left( \sum_{k=1}^{n} \frac{v_k - v_{k-1}}{v - \bar{v}} \mathbb{E}(v|k) \right)^2 > -8(2 - \alpha) \sum_{k=1}^{n} \frac{v_k - v_{k-1}}{v - \bar{v}} (\mathbb{E}(v|k))^2
$$

which is true by concavity of $f(x) = -8(2 - \alpha)x^2$.

Finally, we need to show that there exist a $q'$ such that $\mathbb{E}(CS^C) > \mathbb{E}(CS^{CP})$ if and only if $q < q'$. Both are quadratic and increasing in $q$, but $\mathbb{E}(CS^{CP})$ is increasing faster. Therefore, when $q$ is large enough, $\mathbb{E}(CS^{CP}) > \mathbb{E}(CS^C)$. Then we need to show that if $q = 0$, $\mathbb{E}(CS^C) > \mathbb{E}(CS^{CP})$. The difference between the two expressions is

$$
\mathbb{E}(CS^C) - \mathbb{E}(CS^{CP}) = \frac{(1 - 2\alpha)^4 + 2(3 - 2\alpha)(1 - 2\alpha)^2(\bar{v} + \bar{v}) + (3 - 2\alpha)^2\bar{v}^2 + (4\alpha^2 - \alpha - 7)\bar{v}\bar{v} + (3 - 2\alpha)^2\bar{v}^2}{2(5 - 2\alpha)^2} > 0
$$

which completes the proof.

Proof of Proposition 7.

For this proof, we introduce the expected full revelation revenue, i.e. the revenue the platform would get under competition if $v$ was observable by the sellers:

$$
\mathbb{E}(\pi^{FR}_p) = \frac{(3 - 2\alpha)\phi (3(5 - 2\alpha)^2q^2 + 16 (12\alpha^2 + \bar{v}^2 - 6\alpha(\bar{v} + \bar{v} + 2) + \bar{v}\bar{v} + 3\bar{v} + \bar{v}^2 + 3\bar{v} + 3))}{12(5 - 2\alpha)^2} \quad (A.4)
$$

It follows from Jensen’s inequality that $\mathbb{E}(\pi^{FR}_p) \geq \mathbb{E}(\pi^R_p) \geq \mathbb{E}(\pi^C_p)$.

First, consider the case $\alpha = 0$. It is possible to show that $\mathbb{E}(\pi^{CP}_p)|_{\alpha=0} > \mathbb{E}(\pi^{FR}_p)|_{\alpha=0}$ given the parameter restrictions given by Assumption A1. Then if $\alpha = 0$ centralized pricing is
the most profitable regime for the platform.

Second, consider the case $\alpha = \frac{1}{2}$. In this case under recommendation $n^*$ goes to infinity. Note that now since in this case both sellers are effectively monopolists over their segments of the market, there is no misalignment of incentives between the platform and the sellers. Therefore, when $\alpha = \frac{1}{2}$, the platform will communicate the realization of $v$ to the sellers perfectly. In other words, $E(\pi_P^{FR} | \alpha = \frac{1}{2}) = E(\pi_P^{FR} | \alpha = 0)$. It is possible to show that $E(\pi_P^{FR} | \alpha = \frac{1}{2}) > E(\pi_P^{CP} | \alpha = \frac{1}{2})$ given Assumption A1. Then when $\alpha = \frac{1}{2}$ the most profitable regime for the platform is price recommendation.

Finally, one can show that $E(\pi_P^{CP} | \alpha = 0) > E(\pi_P^{FR} | \alpha = 0)$ given Assumption A1. This means that in the first stage of the game, the platform will choose $\alpha = 0$ and then it will choose to centralize pricing.

A.2. Signaling Model of Buyer Demand

In this section we provide an example of a model that would generate behavior similar to the one presented in the paper without the assumption that the buyers observe the sellers’ qualities $q_1$ and $q_2$. The key insight is that under competition and recommendation, the sellers can use prices to signal their quality while under centralized pricing they cannot.

The setup of the model is the same as before, except for two changes. First, the buyers do not observe $q_1$ and $q_2$ before making purchase decisions. Instead, they form expectations of quality given the information they observe (i.e., the prices). Second, we assume that some buyers can become disgruntled when their expectations are not met and ask for (and receive) a refund with probability $\gamma(E(q_i | p_i) - q_i)$. The probability of refund is a function of the difference between the expected and realized quality. We assume that $\gamma(x) = 0$ if $x < 0$, i.e., no refunds happen after positive surprises. Note that this means that since $q$ is the highest possible quality level, the high quality type never has to pay out a refund.

If $x > 0$, i.e., if the surprise is negative, we assume that $\gamma(x) = \frac{x}{2q}$. Then the probability
of refund is linear and increasing in the size of the negative surprise. Because $\gamma(2q) = 1$, if buyers believe the seller to be of high quality, but it turns out to be low quality, they will always get a refund. Suppose buyers believe that a seller setting price $p$ is of high type with probability $\beta(p)$. Then $E(q_i|p) = (2\beta(p) - 1)q$ and if $q_i = -q$, $\gamma = \beta(p)$. This can be interpreted as whenever a buyer is fooled into buying from a low quality seller, they will ask for a refund.

Given this setup, we can find the following:

**Proposition 8.** Under competition and recommendation there is a separating equilibrium in which all players’ strategies and payoffs are identical to the ones described in Propositions 2 and 3. This equilibrium survives the intuitive criterion refinement.

Under centralized pricing, in the unique equilibrium all players’ strategies and high type sellers’ payoffs are identical to Proposition 1 and the low type sellers’ payoffs are divided by 2 and the platform’s expected revenue multiplied by $\frac{3}{4}$.

**Proof.** We start with competition. Since the strategies are the same as in Proposition 2, we only need to describe the buyers’ beliefs to characterize this equilibrium. The beliefs are the following: $\beta(p_H^C) = 1$ and $\beta(p) = 0 \forall p \neq p_H^C$.

First, we verify that there are no deviations. Suppose the high type deviates. Then it is believed to by low type with probability 1. In this case and given that the other seller is playing the equilibrium prices, the best possible deviation is to $p_H^C$ yielding $E(\pi_L)$, which is less than $E(\pi_H)$, therefore, there are no profitable deviations for the high type.

Now suppose the low type deviates. If they deviate to $p_H^C$, they will earn 0 because of the buyers’ refunds. If they deviate to any other price, they do not shift the buyers’ beliefs and we already know by revealed preference that the optimal price under the belief that the player is low type is $p_L^C$, so there are no profitable deviations for low type. This means that this is in fact an equilibrium.
Second, we show that this equilibrium survives the Intuitive criterion refinement (Cho and Kreps, 1987). This refinement says (in terms of our model) that if for a certain price \( p, \beta(p) > 0 (\beta(p) < 1) \), then it has to be the case that this price is not equilibrium-dominated for high (low type). A price is said to be equilibrium-dominated for type \( \tau \) if 
\[
\max_{\beta} \mathbb{E}(\pi_{\tau}^{\text{dev}}(p, \tilde{\beta})) \leq \mathbb{E}(\pi_{\tau}^C).
\]
In other words, if a price \( p \) is believed to be set with a positive probability by a type \( \tau \) seller, it has to be rationalizable in the sense that there exists a belief \( \tilde{\beta}(p) \), for which this price would be a profitable deviation. Since the original Intuitive Criterion is formulated for a single sender environment, we have to adjust our refinement to assume that for all conditions the other sender’s equilibrium actions are taken as given. This is perhaps the most restrictive version of the refinement.

For a high type seller, for any price, the belief that maximizes their expected profit is \( \tilde{\beta}(p) = 1 \). The best price under this belief is \( p_H^C \). Therefore, for the high type seller, every \( p \neq p_H^C \) is equilibrium-dominated. The out-of-equilibrium beliefs we described above satisfy that.

For a low type seller, if there exists a price that is equilibrium-dominated, then it is equilibrium-dominated for both types and (i) the Intuitive Criterion does not restrict beliefs about these prices (ii) these prices are irrelevant to whether an equilibrium exists or not, since they are not feasible deviations. This means that the out-of-equilibrium beliefs described above satisfy the Intuitive Criterion.

Finally, under centralized pricing, since the platform sets the prices without observing \( q_1 \) and \( q_2 \) and the sellers take no action, there is no opportunity for signalling. The beliefs are \( \beta(p) = \frac{1}{2} \) \( \forall p \), since the buyers know that the platform does not know anything about the sellers’ qualities. This implies that for the low type seller \( \gamma = \frac{1}{2} \), i.e., half of the sales are refunded. For the platform, \( \frac{2}{3} \) of all sales are refunded. As a multiplier, this adjustment does not change the optimal pricing of the platform.

Following this model, the results of the paper might change for specific parameter values,
but will not change qualitatively.
APPENDIX for Chapter 2

B.1. Demand Validation

B.1.1. Validation procedure details

I start by selecting the markets in which an entry occurs. Unsurprisingly, the markets which I selected have a higher median number of licenses (6) than the number of licenses in the markets in the data (5). They also tend to be earlier in the sample period, as later all licenses are taken up and no more entry can occur. Since some retailers enter in the middle of the month, their sales may be artificially low. To avoid this problem, I predict the equilibria in the market for the three months post-entry, excluding the month of entry. To predict the demand and costs for the entrant, I draw from the set of types of the incumbents in the given market. I treat the demand and cost types as correlated. Note that even though the entrant and one of the incumbents in this case are going to have identical demand and cost types, their behavior will still differ as their locations are different. The procedure is the following:

1. pick an incumbent retailer \( r_{inc} \)

2. assign to the entrant \( r_{inc} \)'s demand and cost types

3. solve for the Bertrand-Nash equilibrium in the market

4. repeat steps 1-3 for all incumbent retailers

5. take the average over the equilibrium outcomes (market shares and revenues)

I remove an outlier market in which the entrant sold only 50 grams in its second month and exited the market.
Table 28: Summary statistics for the prediction error in prices

<table>
<thead>
<tr>
<th></th>
<th>individual</th>
<th>entrants</th>
<th>incumbents</th>
<th>market</th>
<th>large market</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>-1.26</td>
<td>-1.18</td>
<td>-1.26</td>
<td>-1.14</td>
<td>-0.54</td>
</tr>
<tr>
<td>5 percentile</td>
<td>-0.40</td>
<td>-0.54</td>
<td>-0.38</td>
<td>-0.56</td>
<td>-0.26</td>
</tr>
<tr>
<td>median</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>95 percentile</td>
<td>0.37</td>
<td>0.66</td>
<td>0.35</td>
<td>0.50</td>
<td>0.23</td>
</tr>
<tr>
<td>max</td>
<td>1.69</td>
<td>1.69</td>
<td>1.58</td>
<td>0.96</td>
<td>0.56</td>
</tr>
<tr>
<td>mean</td>
<td>-0.03</td>
<td>0.01</td>
<td>-0.04</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>st dev</td>
<td>0.26</td>
<td>0.37</td>
<td>0.25</td>
<td>0.29</td>
<td>0.16</td>
</tr>
<tr>
<td>N</td>
<td>2694</td>
<td>264</td>
<td>2430</td>
<td>286</td>
<td>165</td>
</tr>
</tbody>
</table>

B.1.2. Results

On market level, especially for large markets, the model performs reasonably well. It is particularly encouraging that the mean and the median for both prices and quantities are close to zero. On individual level, as expected, there is much more variation. For prices, the error mean and median are close to zero, but the range of values is very wide. (The model does predict negative prices for some retailers and markets, which needs to be fixed in the future). In terms of quantity, the individual predictions for entrants are severely overestimated. Similarly to prices, the model predicts a wide range of values, but market shares cannot go below zero, so the large errors cannot be matched by small ones.

Consider the histograms of errors in Figures 17 and 18. Note that with the exception of the quantity error for entrants, the tails of the distributions are quite thin, which we can also see from large differences between max/min and 95th/5th percentile variables. Price errors have a much more symmetric distribution (even if I truncated the small number of negative prices, the distribution would remain quite symmetric and centered around zero).

In figure 19 I plot the market-level errors in quantity and price with the number of firms in the market. While the mean error does not seem to vary very much with the number of firms (it is always close to zero), the variation in the outcomes decreases dramatically in the number of firms.
Figure 17: Distribution of prediction errors for individual retailers
Figure 18: Distribution of prediction errors on market level

Figure 19: Market-level prediction errors vs number of firms in the market
There could be two reasons why there is less dispersion in errors for larger markets. First, since there are more incumbent retailers, the large error coming from the entrant matters less. Second, the predictions for the entrants might be better in larger markets, as I take more type draws. In figure 20 I plot how the entrants’ errors in quantities vary with the number of firms. To the extent that there is any pattern, it looks like the error increases with the number of firms. This implies that having more draws does not improve the prediction and the smaller error for large markets stems from having more incumbents.
C.1. Robustness Check: Matching only with movies

One concern is that other products cannot be a good control group for movies, since films are a very particular product. To address this concern, we perform the matching exercise and regression analysis using only the set of movies not affiliated with Comcast, Disney or News Corp as a potential control group. We use the same set of cable networks for matching. To control for timing, I consider movies released during the same quarter as the focal movie. I select only 10 matched films for each of the 196 movies of interest, as there are not as many films as other products. The results are summarized in Tables 29, 30 and 31. For NBC-Universal, the results are the same: no significant affiliation effect. For ABC-Disney, the estimated affiliation effects are significant, but smaller in magnitude. For FOX-20th Century, we find no significant affiliation effect if we include Fox Searchlight movies in the sample. However, if we consider only movies released by 20th Century Fox Pictures, the affiliation effect is significant, if smaller in magnitude than the one we find in the main regressions in Table 18. Overall, the main findings of the paper are robust.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>NBC Universal</td>
<td>0.014***</td>
<td>0.011***</td>
<td>0.014***</td>
<td>0.011***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBCXUniversal</td>
<td>-0.010**</td>
<td>-0.013**</td>
<td>-0.006**</td>
<td>-0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.006)</td>
<td>(0.004)</td>
<td>(0.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nobs</td>
<td>858</td>
<td>3432</td>
<td>1716</td>
<td>616</td>
<td>2464</td>
<td>1232</td>
</tr>
<tr>
<td>Nparams</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>R2adj</td>
<td>-0.001</td>
<td>0.015</td>
<td>0.013</td>
<td>-0.002</td>
<td>0.012</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Table 29: Regression results for Comcast movies, using only other films as the set of possible matches.
<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>0.017***</td>
<td>0.018***</td>
<td>0.016***</td>
<td>0.016***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disney</td>
<td>0.011*</td>
<td>-0.012</td>
<td>0.008</td>
<td>-0.011</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
<td>(0.006)</td>
<td>(0.010)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABCXDIsney</td>
<td>0.051***</td>
<td>0.040***</td>
<td>0.063***</td>
<td>0.056***</td>
<td>0.048***</td>
<td>0.067***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.011)</td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.011)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Nobs</td>
<td>484</td>
<td>1936</td>
<td>968</td>
<td>418</td>
<td>1672</td>
<td>836</td>
</tr>
<tr>
<td>Nparams</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.040</td>
<td>0.036</td>
<td>0.065</td>
<td>0.047</td>
<td>0.041</td>
<td>0.065</td>
</tr>
</tbody>
</table>

Table 30: Regression results for Walt Disney movies, using only other films as the set of possible matches.

<table>
<thead>
<tr>
<th></th>
<th>OLS1</th>
<th>DID1</th>
<th>DID2</th>
<th>OLS2</th>
<th>DID3</th>
<th>DID4</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOX</td>
<td>-0.026***</td>
<td>-0.013***</td>
<td>-0.022***</td>
<td>-0.010***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20thCentury</td>
<td>0.005</td>
<td>-0.002</td>
<td>-0.016**</td>
<td>-0.024***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.007)</td>
<td>(0.005)</td>
<td>(0.008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOXX20thCentury</td>
<td>0.010*</td>
<td>0.005</td>
<td>0.012</td>
<td>0.023***</td>
<td>0.039***</td>
<td>0.047***</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
<td>(0.009)</td>
<td>(0.008)</td>
<td>(0.011)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Nobs</td>
<td>814</td>
<td>3256</td>
<td>1628</td>
<td>561</td>
<td>2244</td>
<td>1122</td>
</tr>
<tr>
<td>Nparams</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>R2adj</td>
<td>0.002</td>
<td>0.029</td>
<td>0.012</td>
<td>0.012</td>
<td>0.020</td>
<td>0.016</td>
</tr>
</tbody>
</table>

Table 31: Regression results for News Corp movies, using only other films as the set of possible matches.


H. Guda and U. Subramanian. Your uber is arriving: Managing on-demand workers through


B. Hollenbeck and K. Uetake. Taxation and market power in the legal marijuana industry. *Available at SSRN 3237729*, 2019.


