Your Browsing History May Cost You: A Framework for Discovering Differential Pricing in Non-Transparent Markets

Aditya Karan*
karan2@illinois.edu
University of Illinois
Urbana-Champaign
USA

Naina Balepur*
nainab2@illinois.edu
University of Illinois
Urbana-Champaign
USA

Hari Sundaram
hs1@illinois.edu
University of Illinois
Urbana-Champaign
USA

ABSTRACT

In many online markets we “shop alone” — there is no way for us to know the prices other consumers paid for the same goods. Could this lack of price transparency lead to differential pricing? To answer this question, we present a generalized framework to audit online markets for differential pricing using automated agents. Consensus is a key idea in our work: for a successful black-box audit, both the experimenter and seller must agree on the agents' attributes. We audit two competitive online travel markets on kayak.com (flight and hotel markets) and construct queries representative of the demand for goods. Crucially, we assume ignorance of the sellers' pricing mechanisms while conducting these audits. We conservatively implement consensus with nine distinct profiles based on behavior, not demographics. We use a structural causal model for price differences and estimate model parameters using Bayesian inference. We can unambiguously show that many sellers (but not all) demonstrate behavior-driven differential pricing. In the flight market, some profiles are nearly 90% more likely to see a worse price than the best performing profile, and nearly 60% more likely in the hotel market. While the control profile (with no browsing history) was on average offered the best prices in the flight market, surprisingly, other profiles outperformed the control in the hotel market. The price difference between any pair of profiles occurring by chance is \$ 0.44 in the flight market and \$ 0.09 for hotels. However, the expected loss of welfare for any profile when compared to the best profile can be as much as \$ 6.00 for flights and \$ 3.00 for hotels (*i.e.,* 15× and 33× the price difference by chance respectively). This illustrates the need for new market designs or policies that encourage more transparent market design to overcome differential pricing practices.

CCS CONCEPTS

• General and reference → Empirical studies; **• Information systems → E-commerce infrastructure**; **• Theory of computation →** *Bayesian analysis.*

*Both authors contributed equally to this research.

FAccT '23, June 12–15, 2023, Chicago, IL, USA

ACM ISBN 979-8-4007-0192-4/23/06…\$15.00
https://doi.org/10.1145/3593013.3594038

ACM Reference Format:
Aditya Karan, Naina Balepur, and Hari Sundaram. 2023. Your Browsing History May Cost You: A Framework for Discovering Differential Pricing in Non-Transparent Markets. In *2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23), June 12–15, 2023, Chicago, IL, USA.* ACM, New York, NY, USA, 19 pages. https://doi.org/10.1145/3593013.3594038

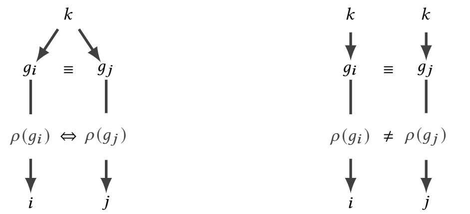
1 INTRODUCTION

What if you found out that you had paid *more* for an airline ticket than your friend who bought an *identical* ticket at the same time? You would be frustrated to learn that the seller charged you a higher price due to your online behavior. This raises the question — does the design of typical online markets encourage differential treatment of buyers? In this work, we show a causal link between behavioral tracking and differential pricing in online markets.

Many online markets lack one key feature: *price transparency*. We say that a market exhibits price transparency if sellers do not withhold prices of previous transactions from consumers. In a transparent market (Figure 1a), buyers can be sure that sellers are offering the same price to different consumers for identical goods at the same time; this price is listed publicly. In an opaque market (Figure 1b), consumers *cannot* be confident that sellers are offering the same price to different consumers for identical goods at the same time; the only price they see is private. Given massive transaction volumes online (nearly 150 million yearly travel bookings [18]); personalized pricing can cause significant loss to consumers *even if sellers employ differential pricing with small probability.*

There are several broad strategies used to study differential treatment. Auditing online markets using automated agents ('sock-puppets') with different constructed attributes is one. Scholars have found disparate treatment in protected markets [6, 17] (the employment and housing markets are protected by federal statutes in the US) and other online markets [25, 26, 28, 31]. In general, these techniques assume that the sellers can correctly infer and act upon the sock-puppets' attributes. A second method is to theoretically model the market [40] and develop policy recommendations based on fairness criteria. But in practice, we often lack follow-up work studying the effects of implementing these fair mechanisms in the real world. Finally, another important approach is to develop seller-pricing algorithms that satisfy fairness criteria [67]. One challenge, however, is that we cannot ensure that all sellers will use these fair pricing algorithms.

In this paper, we conduct sock-puppet audits of two non-transparent online markets to discover differential pricing. We assume ignorance of the sellers' pricing mechanisms, and construct queries representative of varied demand for goods in these markets. Consensus is a key idea in our work: for a successful black-box



(a): transparent price **(b): opaque price**

Figure 1: In a transparent market (a), a seller k offers equivalent goods (g_i, g_j) to buyers (i, j) at *equal* prices $(\rho(g_i) = \rho(g_j))$. Transparent markets can be online or in-person (*e.g.,* the online stock exchange or an in-person grocery store). In an opaque market (b), a seller k offers equivalent goods (g_i, g_j) to buyers (i, j) at *unequal* prices $(\rho(g_i) \neq \rho(g_j))$. Opaque markets too can be online or in-person (*e.g.,* an online flight market or in-person housing market).

audit, we need to show that *both* the experimenter (auditor) and seller agree on buyer attributes. Having established consensus, we can ascribe the seller's change in price between two buyers for the same good at the same time to *only* the difference in attributes. We implement consensus in our audit with nine distinct profiles based on behavior, *not* demographics. We use a structural causal model for price differences and estimate model parameters using Bayesian inference.

We audit a large online flight market $\mathcal{M}_F$, and a large online hotel market $\mathcal{M}_H$ (both at kayak.com) with nine distinct profiles. In $\mathcal{M}_F$, we execute 112 distinct queries every day for 58 days, collecting 1.1M records resulting in 4.2M price comparisons. In $\mathcal{M}_H$, we execute 50 distinct queries over 116 days, collecting 3.6M records resulting in 11.1M price comparisons. In $\mathcal{M}_F$ and $\mathcal{M}_H$, we can unambiguously show that many sellers (but not all) demonstrate behavior-driven differential pricing. The probability that *any* pair of profiles will see a price difference by chance in $\mathcal{M}_F$ is 8.3%; in $\mathcal{M}_H$ it is 2.3%. However, for some profiles and some sellers in $\mathcal{M}_F$, the probability of a price difference *when compared to* the best performing profile can be as high as 90%; in $\mathcal{M}_H$ it can be high as 60%. The price difference between *any* pair of profiles occurring by chance is \$ 0.44 in $\mathcal{M}_F$ and \$ 0.09 in $\mathcal{M}_H$. However, the expected increase in price for a profile *in comparison to* the best performing profile can range as high as \$ 6.00 in $\mathcal{M}_F$ and \$ 3.00 in $\mathcal{M}_H$. In non-transparent markets with millions of transactions, this is a significant impact. Our contributions are as follows:

Generalizable Auditing with Consensus: Our insight is that *explicitly* establishing experimenter-seller consensus regarding buyer identity is critical to conducting audits. In contrast, past work on ad-delivery [6, 17], and price [25, 26, 28, 31, 38] *implicitly* assumes consensus — that the party audited can successfully infer the attributes of the agent set by the researcher, and uses those attributes in its response. We conservatively establish consensus via distinct behaviors, making no assumptions about demographic inferences made by the seller. Our framework for auditing non-transparent online markets generalizes to other online markets.

Causal Differential Treatment: We show significant causal, seller-dependent differential treatment of our profiles in two markets. In contrast, past work [62] could not establish differential treatment in travel prices due to confounds. We develop a structural causal model [51] to identify causal factors for price difference and account for common confounds. We infer model parameters through a Bayesian framework (novel in auditing). In online flight ticket and hotel booking markets we find differential pricing dependent on behavior-based profiles. Opacity of these markets enables this outcome — transparent market design is needed to overcome differential pricing. [1]

2 RELATED WORK

In our work we focus on identifying differential pricing on the basis of behavior. Though we do not aim to audit for bias against any specific demographic group, we do draw from the methods of works that have done so. We utilize prior work to inform our audit methodology and structural causal model.

Auditing Markets Protected by Statutes: In the United States, the Civil Rights Acts of 1964 and 1968 offer protection against employment and housing discrimination on the bases of race, religion, national origin, and sex [13, 42]. Focusing on these markets, both government [44] and private organizations have conducted studies to determine whether lenders [54, 68] or employers [7, 21] are violating these statutes. In online settings, Datta et al. [17] and Asplund et al. [6] use sock-puppet audits to find profile-based discrimination in employment and housing ads respectively. Other works measure bias present in hiring and career recommendation algorithms and propose fixes — usually via algorithms which satisfy some definition of fairness [10, 23, 29, 32, 52, 65].

Auditing Unprotected Online Markets: In other markets, no U.S. federal law explicitly forbids price discrimination based on protected attributes. In such markets, similar products (*e.g.,* deodorant, disposable razors) may be priced differently depending on the target gender, without legal repercussions [12]. Note that in this case the differential pricing is due to marketing or segmentation of the product, rather than being dependent on the buyer's identity (though the product and buyer are likely correlated). Hannak et al. [25] audit various e-commerce websites by collecting the cookies of real-world users. This provides more ecological validity than constructed profiles in typical sock-puppet audits, but still does not guarantee consensus — we do not know what the seller infers. Similar work was done by Hindermann [26], Iordanou et al. [31]. Xu et al. [67] recognize that price personalization on e-commerce websites exists, and propose fairer personalization policies. However, we cannot guarantee that sellers will utilize these fairer policies. Vissers et al. [62] claim that differences in airline ticket prices could come from several hard-to-discern factors. Our sock-puppets and paired-comparison model addresses some of the difficulties mentioned. Mikians et al. [38] train personas based on personal (*e.g.,* affluent or budget-conscious), geographic, and technological/system

[1] Code to train agents and model data can be found at https://github.com/CrowdDynamicsLab/browsing-history-cost

information. They then query websites for different products and find price discrimination on these profile attributes.

Auditing ML Systems: The first major works on detecting and mitigating algorithmic bias focused on the resulting policies and major social implications of biased algorithms [4, 48]. Costa-jussà et al. [14], Matthews et al. [37], Park et al. [50] all propose methodologies for assessing bias in machine learning: specifically gender bias in machine learning, fairness and security of machine learning algorithms, and bias in word embeddings. Feng and Shah [22] investigate whether search engines properly mitigate gender bias, and propose new fair ranking algorithms that could defend against adversarial attacks, along with Tsioutsiouliklis et al. [60]. These audits for bias have grown substantially over the past several years [15]. Some are done in conjunction with system providers [65], which can lead to charges of conflict of interest [69]. Crucially many of these audits [1] require using explicit input features which may not be available in all settings.

Ad-Tracking Ecosystem: Researchers have also investigated the connection between web browsing and the ad ecosystem itself. Understanding the link between browsing behavior and personalization online is vital to the validity of our experimental design. Angwin [3], Englehardt et al. [20], Sipior et al. [58] study the implications of web tracking for users. Prior work has attempted to determine how browsing behavior and user attributes covary [34], however covariation does not guarantee consensus even if found.

Price Discrimination: Price discrimination in economics refers to the effort of firms to selectively charge based on the buyer's willingness to pay for a specific product [39, 61]. The ability of firms to engage in price discrimination depends heavily on the competitiveness of the market [59]. Note that willingness to pay can depend on the opacity of the market; most consumers dislike when prices are personalized and may not pay more if they know another consumer is being offered a lower price [70]. Previous works try to determine how differential pricing affects consumers. Dubé and Misra [19] find theoretically and empirically that the total consumer surplus decreased under personalized pricing compared to uniform pricing; however, a majority of consumers actually ended up better off. Khan [33] explores the rise of Amazon via predatory pricing behaviors. There has also been related work on discrimination unrelated to price. The theoretical model by Monachou and Ashlagi [40] describes biased hiring behavior — employers discriminate against workers based on their (perceived) social status. They find that minority workers receive lower payoffs and propose a matching strategy that decreases discrimination. Cui et al. [16] discover that positive reviews on Airbnb decrease discrimination against guests with African-American sounding names.

Transparency in Markets: In seminal work, Bloomfield and O'Hara [9] find that varying market transparency can affect market equilibria and both trader and market-maker welfare. Scalia [57] conducts a case study on Italian bonds, and finds that decreasing transparency leads to higher liquidity and lower volatility. Ionascu et al. [30] study how market transparency affects housing prices, and find that a decrease in transparency is related to an increase in prices. Researchers also consider market design in general; Levy and Barocas [35] consider ten categories of design and policy choices to determine which may decrease discrimination in online markets. Armstrong [5] finds that whether consumers overall are better off depends heavily on market structure.

3 PROBLEM STATEMENT

Consider a large online market $\mathcal{M}$ for indivisible goods G, with a set of buyers N and sellers S. Each buyer $i \in N$ has attributes $a_i \in A$ (*e.g.,* age, behaviors); we refer to a_i as the profile of individual i. Buyer i issues a query $q \in Q$, where Q is the set of all valid queries in this market. A query q is defined both by the query text and the time at which it was executed (*i.e.,* two queries for "blue t-shirts" executed at *different* times are *not* identical). The set of goods relevant to this query is denoted G^q. Each seller $k \in S$ is only able to sell a subset of these goods, and of these, they offer a subset to each buyer i. The market $\mathcal{M}$ aggregates the responses of all sellers and reveals to each buyer $i \in N$, the set of goods relevant to the query: G_i^q.

Each good $g \in G$ has attributes b_g, a vector of product details. The price of the good in dollars is notated $\rho(g)$. Assume two buyers $i, j \in N$ issue identical queries q at the same time. In response, assume that seller $k \in S$ reveals to buyers i, j, goods $g_i \in G_i^q$ and $g_j \in G_j^q$ respectively. We say that the goods g_i, g_j are *equivalent* (*i.e.,* $g_i \equiv g_j$) if two conditions hold: (1) the two goods have identical product details, (*i.e.,* $b_{g_i} = b_{g_j}$), and (2) g_i and g_j are sold by the same seller k at the same time. We say a market lacks price transparency if individuals i and j cannot be sure that $\rho(g_i) = \rho(g_j)$, when $g_i \equiv g_j$.

Assume that two buyers i and j with attributes $a_i \neq a_j$ issue identical queries q (*i.e.,* at the same time), in market $\mathcal{M}$ with sellers S. For the following questions, assume that good $g_i \in G_i^q$, good $g_j \in G_j^q$, and $g_i \equiv g_j$. We ask:

RQ 1 (*Differential Treatment*) For all buyers $i, j \in N$ what is the probability that i and j are offered different prices for equivalent goods, by the same seller, at the same time?

RQ 2 (*Differential Loss*) For all buyers $i, j \in N$ what is the expected dollar amount lost?

Our research questions are significant. Differential pricing could cause major welfare loss to consumers, but most online markets do not have laws regulating this. Moreover, much of the existing work on differential pricing [6, 25] focuses on steering of goods (*i.e.,* where $b_{g_i} \neq b_{g_j}$) while we focus on price differences between equivalent goods.

4 THE CONSENSUS CHALLENGE

Establishing experimenter-seller consensus on the buyer's attributes is critical to a successful black-box audit. To test for the existence of differential treatment, **RQ 1** (*i.e.,* $\mathbb{P}(\rho(g_i) \neq \rho(g_j))$), we must be able to experimentally manipulate attributes a_i and a_j of individuals i and j respectively. Moreover, we must ensure that the seller perceives the same difference in attributes. Audits of online and physical markets differ in their abilities to guarantee consensus.

Consider a typical physical audit conducted by the U.S. Department of Housing and Urban Development (HUD); two individuals, differing only by a protected attribute (*e.g.,* a young White male and young Black male) contact brokers and observe which housing options and prices are shown to them [64]. When constructing these audits, the auditor ensures that the broker perceives a difference in the race (not gender or age) of the buyers, either by physical appearance or racially-coded names [47].

In many online marketplaces, sellers do not receive clear signals regarding the identities of the buyers. Early work on bias in employment ads [17] used automated agents with explicitly set attributes via Google Ad settings. Google and other companies now disallow this practice. Can we successfully manipulate attributes without direct access?

As Papadopoulos et al. [49] carefully document, a large number of corporations track individuals' movements over the web, inferring detailed behavioral profiles for each person, including demographic information (*e.g.,* gender, race, age) and interests. However, the process of inferring individual attributes from website visits is not well-understood, and tracking companies may differ in their inference methods. Furthermore, since sellers may use inferences made by different tracking companies, sellers may make *different* inferences about the *same* behavior.

We anchor the notion of consensus in this work to an individual i's online behavior: the *set of websites* visited, $w_i \in \mathcal{W}$. That is, we set $a_i = w_i$, where w_i is the set of websites visited by individual i. In our work, a_i refers to behavior, and *not* demographic attributes. We make two conservative assumptions: (a) individuals are tracked and (b) both seller and experimenter agree that the individuals behaved distinctly; this is key to our model. We consider the behavior of two profiles distinct if they visited different sets of websites. We make no assumptions about the technology used for tracking, *who* tracks the individual, or the inferences made by the trackers. Findings by Papadopoulos et al. [49] support our assumptions. The authors state that we can find trackers on almost all websites, with Google and Facebook trackers found on over 90% of websites. Our own investigation of website trackers also supports this claim.

5 EXPERIMENTAL DESIGN

In this section we discuss which markets we audited (Section 5.1), how we constructed behavioral profiles (Section 5.2), and the data collection process (Section 5.3).

5.1 Markets Audited

We required online marketplaces with the following properties for our audits: the market offers a large number of indivisible goods; goods are offered by multiple sellers; goods are uniquely identifiable by the same properties; goods are consistently offered; the market is liquid; an account is not required to purchase goods. Some well-known online markets do not have these desired properties. For example, though amazon.com offers a large number and variety of indivisible goods, there may not be multiple sellers for each good, creating a confound in our analysis.

This led us to examine two markets, online flight tickets ($\mathcal{M}_F$) and online hotel bookings ($\mathcal{M}_H$), via an aggregator (kayak.com). In these markets, each good may be available for sale by multiple different sellers. In $\mathcal{M}_F$ we observed 34 sellers; in $\mathcal{M}_H$ there were 427. We anonymize sellers in each market, labelling them "First Party" if they operate flights or own hotels, and "Third Party" otherwise. Our numbering convention for flight sellers is 1-indexed and 1000-indexed for hotel sellers.

5.2 Behavioral Profile Construction

Our goal in training profiles was to create *distinct* profiles that *help establish consensus.* We set the web-browsing behavior of the agent $w_j \in \mathcal{W}$ to be the defining attribute a_j. We obtain these website lists w_j from existing literature [2, 6]. Agarwal et al. [2] conducted work on hyper-partisan tracking, and constructed website lists for profiles based on the demographics of visitors to these websites as reported by alexa.com from October 2018 (Alexa has since been retired). In order to create relatively compact profiles, the authors used the amount of unique cookie information as a stopping criterion — *i.e.,* when the number of unique cookies stabilizes, the profile construction is complete. Asplund et al. [6] conducted an online housing market audit; to construct profiles, they collected websites' demographic information from Quantcast (quantcast.com). They added a website to a particular demographic profile if that website was more that 1.4 times more popular than the average website within that demographic, while being of average popularity for other groups.

Prior work associates demographic (gender and age) attributes with these sets of websites. Asplund et al. [6] tried to validate the identity of their profiles indirectly using correlation, but this was insufficient to guarantee consensus due to unobserved confounds. We do not make assumptions about demographic attributes because we cannot guarantee that our inferences match those of the seller. The resulting profiles from prior work are male ($\mathrm{M_A}, \mathrm{M_B}$), female ($\mathrm{F_A}, \mathrm{F_B}$), youth ($\mathrm{Y_A}, \mathrm{Y_B}$), and senior ($\mathrm{S_A}, \mathrm{S_B}$) profiles from [2] and [6] respectively, as well as a control profile (no websites visited — equivalent to incognito browsing). We utilize initials (*e.g.,* $\mathrm{M_A}$) in this paper to de-emphasize identity while allowing readers to make correspondences with original cited work.

5.3 Data Collection

Input : market $\mathcal{M}$, queries Q, agents $\mathcal{N}$ (each with website lists $w_i \in W$), query days D
Output : Records R

for $t \in D$ **do**
  for $q \in Q(t)$ **do**
    for $i \in \mathcal{N}$ **do**
      i visits websites $w_i \in W$
    end
    In parallel processes, every agent queries market $\mathcal{M}$ with q
    Market shows agent i goods g as record r_i
    $R \leftarrow R \cup (\cup_i r_i)$
  end
end

Algorithm 1: Data Collection Process

We collect data by building off code by Datta et al. [17]; see Algorithm 1. Specifically, we visit the website list w_i for agent i

each day prior to executing each query. For the daily website visits we use a new browsing session for each agent with no previous cookie information carried over. This ensures that the agents differ only in browsing behavior. We executed queries for all flight routes on July 6 – July 31, Oct 14 – Nov 8, and Dec 4 – Dec 9, 2021. An example flight query is: "find roundtrip tickets from Los Angeles (LAX) to Chicago (ORD) departing June 11, 2023 and returning June 16, 2023." The system then aggregates relevant results, of which we collected the first page (~20 per page). For hotels, we looked at stays from Oct 4, 2022 – Jan 27, 2023. An example a hotel stay query is: "find one hotel room in Chicago from June 11, 2023 to June 16, 2023 for one guest." The aggregator would again offer relevant results, from which we collected the first 5 results for each star classification of hotel (5★, 4★, etc.). For each hotel shown, we collected every room configuration (*e.g.,* king bed, queen bed with breakfast included, etc.). From $\mathcal{M}_F$ we collected 1.1M records from 34 sellers (9 first party and 25 third party), where the average ticket price was \$ 270.45. From $\mathcal{M}_H$ we collected 3.6M records from 427 sellers (23 third party and the rest first party) where the average room cost was \$ 271.99. As described in Section 3, we compare equivalent goods from the same seller to detect differential treatment of profiles. After processing our dataset to find the equivalent goods from the same sellers, we had 4.2M and 11.1M unique price comparisons between profiles for flights and hotels respectively. Further details are in Appendix C.

6 MODELING DIFFERENTIAL TREATMENT AND PRICE DIFFERENCES

In this section we give details of our structural causal model (Section 6.1), main model parameter estimation (Section 6.2), and model for price differences (Section 6.3).

6.1 Structural Causal Model

In this audit, we experimentally control profiles $a_i \in A$ and queries $q \in Q$. To eliminate confounds which may influence the price of a good, we model the effect of profile variation on price differences, rather than the effect of profile on price. We use the structural causal model (SCM) framework (introduced by Wright [66] but more recently popularized by Pearl [51]) to identify our price difference model. An SCM encodes causal relations between factors that the modeler believes could affect the outcome. In these causal diagrams (DAGs), $x \rightarrow y$ means x causes y.

Consider this causal model for ticket price (P) in Figure 2a. We experimentally control A and Q, hence no arrow. A query Q to the market causes the market to aggregate seller responses for the buyer. Thus we expect the query Q to cause the seller K to appear (*i.e.,* $Q \rightarrow K$). For a buyer to see good G, both the query and the seller must cause it (*i.e.,* $Q \rightarrow G \leftarrow K$) to appear. The seller K, the good G's attributes, and unobserved market confounds U_2 (*e.g.,* supply and demand) may all affect the price (P) directly. If the seller knows the identity of the buyer (*i.e.,* the profile A), then their interaction (Z) may affect the price (*i.e.,* $Z \rightarrow P$). The profile (A) cannot affect price directly, but may through an unobserved confound U_1 (*e.g.,* the market $\mathcal{M}$ may track the buyer independent of any seller and influence the results).

Notice two key challenges: a seller K may have a sophisticated pricing algorithm *independent* of the buyer, and unknown to us. Second, unobserved confounds like supply and demand (U_2) may obscure the effect of changing the profile on price. We cannot estimate differential treatment by modeling price directly, because we cannot estimate these effects on price.

The DAG for price difference eliminates unobserved confounds. We answer **RQ 1** by examining if changing the profile from A_i to A_j causes a change in the price ($P_i \neq P_j$?), while conditioning on seller, query, and good. Figure 2b is the DAG for prices shown to two different profiles A_i, A_j by the same seller K, conditioned on query Q for a particular good G. Notice that the seller (K), the good attributes (G), and the unobserved external factors (U_2) are all directly affecting *both* prices P_i and P_j. We show the direct effects on the two profiles with different colors for clarity. Because of their *common* direct effect on each price (for the *same* good G), the seller K, the good attributes (G) and the external effect U_2 cannot cause any price *difference* $P_i \neq P_j$. A seller K can indirectly affect prices through the interaction Z_i, Z_j with the profiles. Unobserved confound U_1 was also not eliminated because it can affect profiles differently, thus causing a price difference. The DAG we use after eliminating these arrows is Figure 2c.

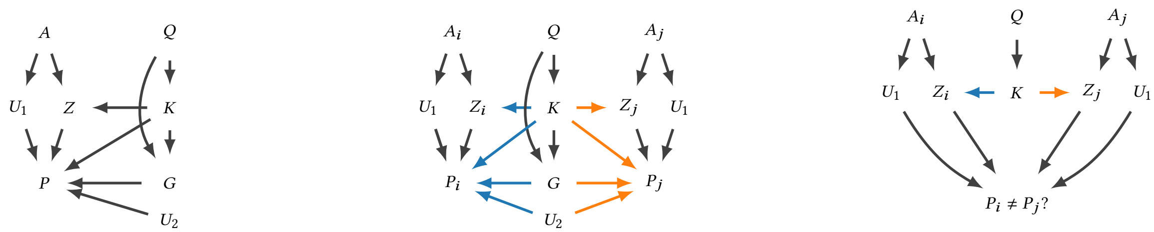
6.2 Model Parameter Estimation

We use Bayesian inference to model the probability that one profile will see a higher price than another for the same seller and good given a query ($y_{i,j|k}$). We model the outcome with a Bernoulli distribution with parameter $p_{i,j|k}$, the probability that buyers i, j will see different prices conditioned on the same seller k (Equation (1)). The outcome $y_{i,j|k} \in \{0, 1\}$ is conditioned on the *same* good g, but we do not include g in the equations for clarity. Notice that outcome $y_{i,j|k}$ is equivalent to asking if $P_i \neq P_j$ in Figure 2c.

$$y_{i,j|k} \sim \mathrm{Ber}(p_{i,j|k}) \quad (1)$$
$$\mathrm{Logit}(p_{i,j|k}) = \bar{p} + (\beta_i - \beta_j) + (\delta_{k,i} - \delta_{k,j}) \quad (2)$$
$$\beta_i \sim \mathrm{N}(\bar{\beta}, \beta_\sigma) \quad (3)$$
$$\bar{\beta} \sim \mathrm{N}(0.0, 1.0) \quad (4)$$
$$\beta_\sigma \sim \mathrm{Exp}(1.0) \quad (5)$$
$$\delta_{k,i} \sim \mathrm{N}(0.0, 0.25) \quad (6)$$
$$\bar{p} \sim \mathrm{N}(-3.0, 1.0) \quad (7)$$

Equation (2) gives the logit link function to estimate the probability $p_{i,j|k}$. Notice similarities to the Bradley-Terry Model [11] which predicts the outcomes of paired comparisons. We can model the effect of the profile a_i through the unobserved confound with parameter β_i. These parameters affect the probability of seeing a different price through the difference in their direct effects, $(\beta_i - \beta_j)$. We capture the seller-profile interaction effect through the parameter $\delta_{k,i}$. As a reminder, this is the only way a seller can influence price difference (see Figure 2c, where Z_i is an interaction variable). This parameter helps us answer to what extent the price difference reflects sellers' strategies. Thus, we introduce $\delta_{k,i} - \delta_{k,j}$, the difference of the interaction effects to understand their effect on the price difference.

Equation (3) shows our multi-level modeling for the direct effects β. We draw β_i from a common prior $\mathrm{N}(\bar{\beta}, \beta_\sigma)$, pooling and



(a): DAG for price (b): DAG for two prices (c): DAG for price difference

Figure 2: The causal graph for price (a) shows the direct effect due to the seller (K), the good (G), and unobserved market conditions (U_2). Price is also affected by the seller-profile interaction (Z). The query Q causes the seller K to respond. The profile (A) can only affect price through an unobserved confound (U_1). Figure (b) shows the DAG for prices for two profiles A_i, A_j for a query Q and good G by the same seller K. We abuse price notation to highlight the price for G shown to the two profiles. Colored arrows highlight effects of K, G, and U_2 on P_i and P_j. Figure (c), the DAG for price difference, eliminates direct effects of seller, good, and market conditions.

regularizing our estimates for β_i. We use weakly informative priors for $\bar{\beta}$, β_σ, and $\delta_{i,k}$ (Equations 4 through 6). We set the mean of the prior for both $\bar{\beta}$ and $\delta_{i,k}$ to be 0, because a priori, we do not expect either to affect the outcome.

Finally, Equation (7) shows the prior for $\bar{p}$, which gives the probability we see a difference in price purely by chance. Our prior belief is that there is no differential treatment in the market. This is why the mean μ of the prior for $\bar{p}$ is set to $\mu = -3$. On the logit scale, this implies that prior probability of seeing a price difference is $\approx 4.7\%$; $\mu = 0$ would imply the prior belief of observing a price difference by chance was 50%. In the case of a single seller market, one can remove Equation (6) and the δ parameters in Equation (2); any effect of that single seller will be subsumed by the β parameters.

We implemented the model in NumPyro (num.pyro.ai) [8, 53] using SVI (Stochastic Variational Inference) [27]. We opt for SVI because the size of our datasets (4.2M observations for $\mathcal{M}_F$ and 11.1M for $\mathcal{M}_H$) makes MCMC too computationally expensive to be feasible. Our original MCMC implementation for $\mathcal{M}_F$ ran for 163 hours on a 10 core, 3.6GHz iMac with an Intel Core i9 processor, and then crashed.

6.3 Model for Price Differences

To answer **RQ 2**: the expected dollar loss that a profile incurs, we create the mixture model depicted in Figure 6 in Appendix B. This model in aggregate has the effect of a zero-inflated Gaussian.

Our idea for this mixture model is to use the probability of a price difference outputted by our original model (1), and multiply this by the expected price different between profiles *when* it exists. This will give us the expected dollar loss of each profile $a_i \in A$ compared to $a_j \in A$. We calculate it as follows, where $Y_{i|k} - Y_{j|k}$ is the expected price difference between profiles a_i and a_j *when* it exists: $\mathbb{E}[\rho(g_i) - \rho(g_j)] = 0 \times (1 - p_{i,j|k}) + (Y_{i|k} - Y_{j|k})(p_{i,j|k}) = (Y_{i|k} - Y_{j|k}) \times (p_{i,j|k})$. Thus, to find the expected loss, we need a model that gives us the dollar difference $Y_{i|k} - Y_{j|k}$ as an output. We already know $(p_{i,j|k})$, as this is the output of our original model.

We choose to tackle the modeling in this way because our price difference data (in dollars) is zero-inflated — most of the time, sellers offer profiles the same price for the same object, resulting in a \$ 0.00 price difference. Though we were able to easily model the zero-inflated data with the Bernoulli distribution (see $\bar{p}$ in Equation (2)), we expected that the price-difference data would result in a Gaussian distribution, and introducing $\bar{p}$ to a Gaussian distribution does not have the same affect of skewing the data toward an outcome of zero. Thus we needed to model the zeros separately from the non-zeros. This model is quite similar to the original, however we only look at a subset of the data. 190k of the rows out of 4.2M for flights and 204k out of 11.1M for hotels in our original datasets have a nonzero price differences — these rows are the input to this model. The differences in the model itself are minimal: we simply model the outcome with a normal distribution rather than Bernoulli, and change some of the prior distributions to match the shift to dollar outcome. This model is fully detailed in Appendix B.

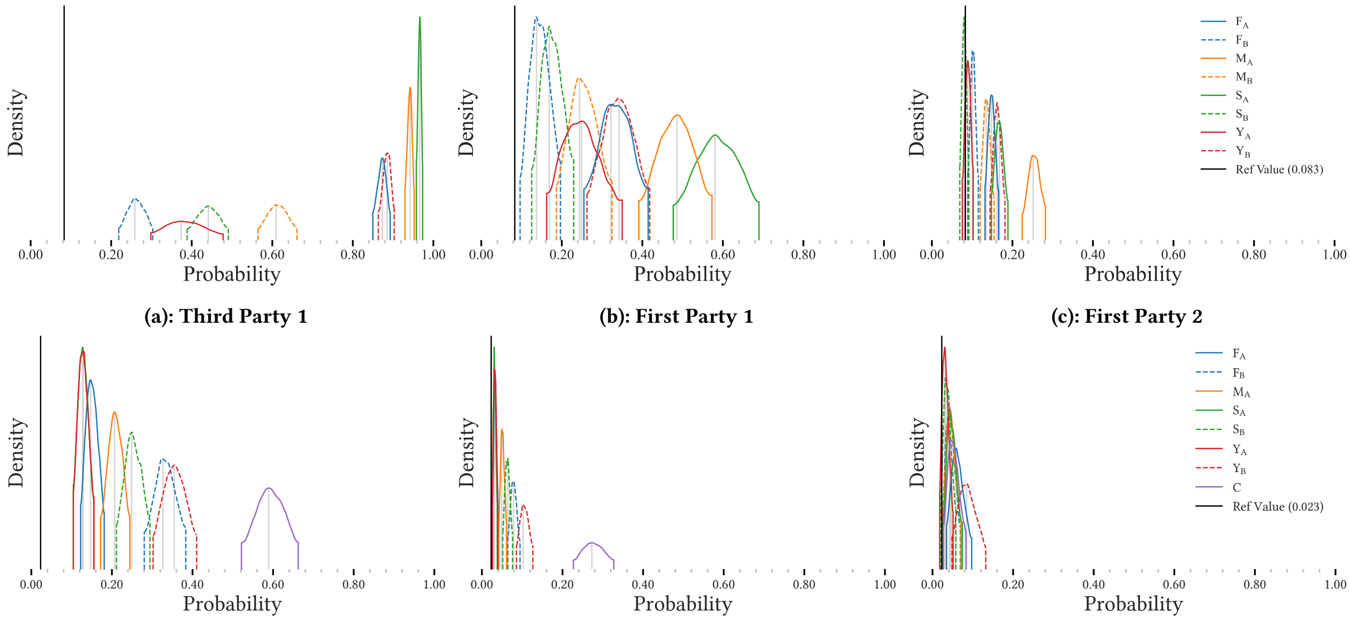
7 RESULTS

In this section we give the results of our model, first to answer **RQ 1** (Section 7.1) then **RQ 2** (Section 7.2).

7.1 RQ 1: Differential Treatment

We observe differential treatment across profiles; the degree is seller dependent. Figure 3 illustrates differential treatment for three select sellers in the flight and hotel markets respectively (full results in Appendix A).

Consider Figure 3a in detail, which gives results for a flight seller. The figure shows for each profile a_i, the posterior probability of receiving a higher price than control (C) for this seller. The black reference line is the probability (8.3%) that *any* pair of profiles see different prices by chance, even if $\beta_i = \beta_c$ and $\delta_{k,i} = \delta_{k,c}$ (see Equation (2)). We show the 89% HDPI region for each density $p_{i,c|k}$. We see that profiles $\mathrm{M_B}, \mathrm{F_B}, \mathrm{Y_B}, \mathrm{S_B}$ receive significantly different prices (*i.e.,* the curves do not overlap). Similarly, all profiles $\mathrm{M_A}, \mathrm{F_A}, \mathrm{Y_A}, \mathrm{S_A}$ differ significantly in terms of price. Some profiles also receive similar treatment; $\mathrm{F_A}$ and $\mathrm{Y_B}$ do not differ significantly in the probability of seeing a worse price than C. Looking at profiles designed to represent the same demographic groups — we observe that $\mathrm{S_A}$ is 96% to 97% more likely to see a higher flight price than C while $\mathrm{S_B}$ is 39% to 49% more likely, with no overlap in the HPDI intervals. The



(a): Third Party 1 **(b): First Party 1** **(c): First Party 2**

(d): Third Party 1015 **(e): Third Party 1016** **(f): First Party 1002**

Figure 3: *Probability of Differential Treatment*: Density plots for $p_{i,j|k}$ (probability of profile a_i seeing a worse price than a given reference a_j). The top row shows a comparison with the control profile in the flight market and the bottom row shows a comparison with $\mathrm{M_B}$ in the hotel market. The plots show 89% HPDI regions; non-overlapping regions imply significant differences in effects. The reference value, a black vertical line, represents the base probability of *any* profile seeing a worse price than any other by chance. Some sellers demonstrate large, distinctive differences: in (a) $\mathrm{M_A}, \mathrm{S_A}$ have a greater than 90% chance of seeing a worse price than control. In (d) we see that the control profile has nearly a 60% chance of seeing a worse price than $\mathrm{M_B}$. Not all sellers engage in much differential pricing between profiles (First Party 2 in (c) and First Party 1002 in (f)).

posterior distribution variance for $p_{i,c|k}$ differs by profile for this seller as well: $\mathrm{S_A}$ has lower variance than $\mathrm{S_B}$.

Now consider Figure 3e, which gives results for a hotel seller. In this figure, we show for each profile $a_i \in A$, the posterior probability of receiving a higher price than $\mathrm{M_B}$. Our reference profile for $\mathcal{M}_H$ is $\mathrm{M_B}$ because this is the profile that received the best treatment from sellers on average. The black reference line here is at 2.3%, representing the probability that two profiles see different prices by chance. We see that the control profile stands apart from others, having between a 52% and 66% chance of seeing a higher hotel price than $\mathrm{M_B}$. Profiles $\mathrm{F_A}, \mathrm{S_A}$ and $\mathrm{Y_A}$ cluster together from 10% to 15% and $\mathrm{Y_B}, \mathrm{F_B}$ are around 28% to 41%. We can also observe that between these two clusters there is a noticeable difference in the width of HDPI intervals.

The differential treatment we detected in these markets can come from two sources: 1) interaction between seller and profile, and 2) factors independent of the seller (*e.g.,* the market itself). First, how much of the differential treatment between profiles a_i and a_j is due to the seller, profile interaction *alone*? To answer this question, we examine the interaction effect variable $\delta_{k,i}$ in Equation (2). We show a subset of the $\delta_{k,i}$ in Figure 4, highlighting sellers with the most and least variation across profiles (full results in Appendix A). Let us revisit the seller Third Party 1 in $\mathcal{M}_F$. Notice that for this seller, $\delta_{k,\mathrm{S_A}} = +1.89$ while $\delta_{k,\mathrm{S_B}} = -0.55$. We now want to calculate the odds of seeing a price increase if we were to change from profile $\mathrm{S_B}$ to $\mathrm{S_A}$. Let $O_{j,i|k}$ be the odds of seeing a price increase when comparing profile a_i to a_j, for seller k. With a little algebra on Equation (2), we show that the *change in* odds for profile a_j is:

$$O_{j,i|k} = \underbrace{\mathrm{Exp}(\beta_j - \beta_i)}_{\text{Non-seller effect}} \times \underbrace{\mathrm{Exp}(\delta_{k,j} - \delta_{k,i})}_{\text{Seller effect}} \quad (8)$$

Thus, when changing from $\mathrm{S_B}$ to $\mathrm{S_A}$, for this seller, the odds of seeing a higher price increased by a factor of $\mathrm{Exp}(\delta_{k,\mathrm{S_A}} - \delta_{k,\mathrm{S_B}}) = 11.47$. $\mathrm{S_A}$ is more than *11 times* more likely to see a worse price than $\mathrm{S_B}$. Conversely, some cases result in little effect on the odds. For seller Third Party 1015 in $\mathcal{M}_H$, we can calculate a similar ratio: $\delta_{k,\mathrm{M_B}} = +1.72$ while $\delta_{k,\mathrm{c}} = -1.57$. Here, the control is 26.84 times more likely ($\mathrm{Exp}(\delta_{k,\mathrm{M_B}} - \delta_{k,\mathrm{c}})$) to see a worse price than $\mathrm{M_B}$.

Next, consider the effect of factors independent of seller on differential treatment. Recall that in Figure 2c we posited an unobserved confound U_1. One possibility is that the market $\mathcal{M}$ mediates (via U_1) the prices by examining the profile. In our model, the coefficients β_i in Equation (2) explain this effect. We show these coefficients in Table 1; all posteriors have sharp HPDI intervals. For the profiles $\mathrm{S_A}$ and $\mathrm{S_B}$ in $\mathcal{M}_F$, the corresponding coefficients are $\beta_{\mathrm{S_A}} = +1.14$ and $\beta_{\mathrm{S_B}} = 0.00$. Thus the effect on the odds due to non-seller effects *alone* is $\mathrm{Exp}(\beta_{\mathrm{S_A}} - \beta_{\mathrm{S_B}}) = 3.12$ (see Equation (8)). Notice, in contrast, had we changed the profiles from $\mathrm{F_B}$ to C in $\mathcal{M}_H$, we would see a *decrease* in odds by a small factor of 0.97.

	C	F_A	F_B	M_A	M_B	S_A	S_B	Y_A	Y_B
Third Party 17	0.00	0.01	-0.01	0.02	-0.00	0.01	-0.01	-0.00	0.02
First Party 4	-0.00	-0.00	0.01	0.01	0.04	0.01	-0.06	-0.00	0.01
First Party 1	-0.42	-0.04	-0.03	0.88	0.32	0.07	-0.67	-0.00	-0.07
Third Party 9	-0.49	0.03	-0.19	0.90	0.23	0.22	-0.71	-0.00	0.01
Third Party 13	-1.59	1.12	-0.50	2.36	0.63	1.85	-0.60	0.29	1.26
Third Party 5	1.19	-0.95	1.19	-0.33	0.69	-2.92	-0.68	-0.42	-0.82
Third Party 22	-1.96	0.76	-0.86	2.05	0.23	1.42	-1.00	-0.42	0.90
Third Party 1	-1.59	1.39	-0.48	2.55	0.68	1.89	-0.55	-0.54	1.45

(a): Flights, $\mathcal{M}_F$

	C	F_A	F_B	M_A	M_B	S_A	S_B	Y_A	Y_B
First Party 1356	-0.00	0.00	-0.01	0.00	-0.00	0.00	-0.01	-0.00	0.00
First Party 1301	0.00	0.01	0.00	0.01	0.01	0.00	0.01	0.01	-0.00
First Party 1338	0.00	-0.00	-0.01	-0.01	0.00	-0.00	0.01	-0.01	-0.00
First Party 1396	0.00	-0.00	-0.00	-0.00	-0.00	0.01	0.01	-0.00	0.00
Third Party 1016	1.89	-0.38	0.36	0.23	-0.06	-0.72	0.34	-0.02	1.04
First Party 1092	0.08	-0.89	-0.83	-0.57	0.91	-0.14	-0.49	0.56	1.33
First Party 1255	0.59	-1.58	-0.46	-0.16	0.23	-0.66	0.75	-0.03	1.23
Third Party 1015	1.57	-0.39	0.46	0.16	-1.72	-0.86	0.29	-0.19	0.92

(b): Hotels, $\mathcal{M}_H$

Figure 4: *Seller effects*: We show the interaction variable $\delta_{k,i}$, for a subset of high and low variance sellers on the logit scale for both $\mathcal{M}_F$ and $\mathcal{M}_H$. We see that some sellers have high variation (Third Party 1) while others have very low (First Party 1356). From this we can see that certain sellers can have nearly opposing strategies: Third Party 1 vs. Third Party 5 in $\mathcal{M}_F$ and Third Party 1015 vs. First Party 1092 in $\mathcal{M}_H$. In $\mathcal{M}_F$ we see that control often does better (but not always) while in $\mathcal{M}_H$ we see the reverse.

These results demonstrate that some sellers do engage in differential pricing on the basis of profile. To understand the full scope, we also need to examine the dollar impact.

Table 1: The mean values of the β_i coefficients in both markets where C is the control profile. The values are in the Logit scale where ± 1 changes in β cause the odds to change by a factor e and e^{-1} respectively (see Equation (8)).

	C	F_A	F_B	M_A	M_B	S_A	S_B	Y_A	Y_B
$\mathcal{M}_F$	-1.12	+0.22	-0.85	-0.08	-0.53	+1.14	+0.00	-0.23	+0.28
$\mathcal{M}_H$	+0.02	-0.13	0.05	-0.28	-0.83	+0.16	-0.17	-0.51	-0.31

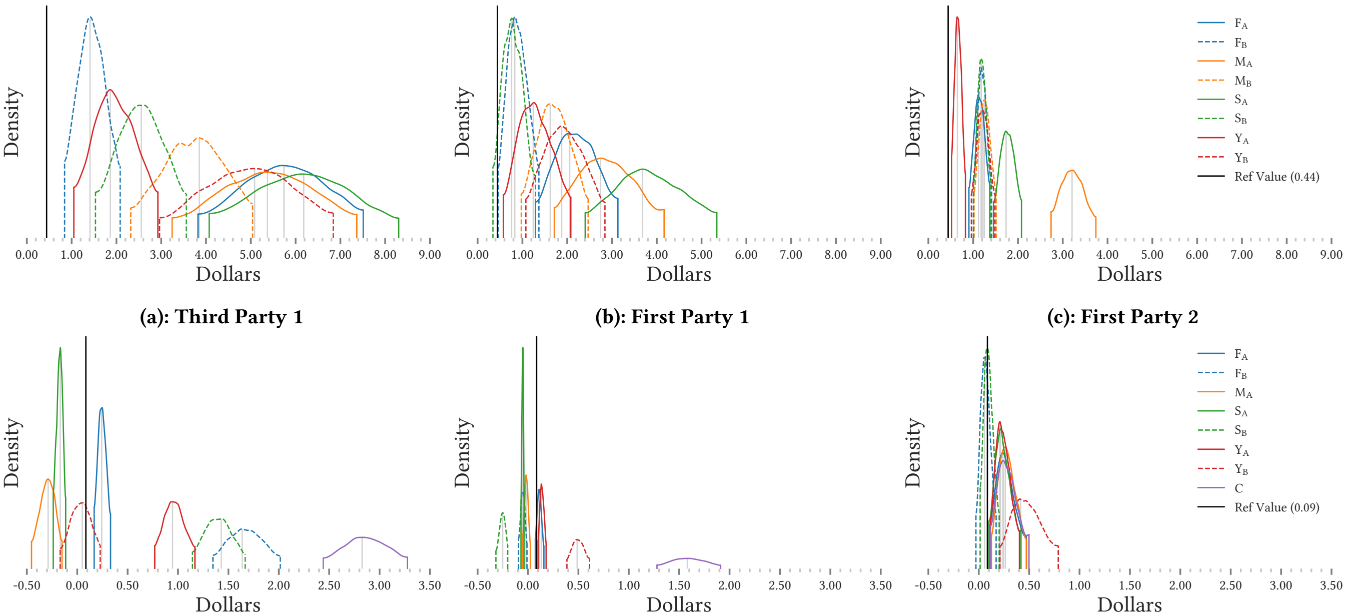
7.2 RQ 2: Expected Differential Loss

Now we examine **RQ 2**, and specifically calculate the expected differential loss $\mathbb{E}(\rho(g_i) - \rho(g_j))$, that a profile i incurs compared to another when $g_i \equiv g_j$,. We use a zero-inflated Gaussian to model price difference $\rho(g_i) - \rho(g_j)$ for profiles a_i, a_j. The data is zero-inflated because profile pairs may see an identical price from the same seller for the same object. Despite possible equal treatment among profiles, they may see a significant price difference with the control profile. In the raw data we observe that in $\mathcal{M}_F$, 25 sellers (out of 34) engage in differential pricing; in $\mathcal{M}_H$ we see this from 134 sellers (out of 427).

We show the results in Figure 5, on dollar scale (full results in Appendix B). The black reference lines are at \$ 0.44 for flights and \$ 0.09 for hotels; they represent the average price difference between any pair of profiles, occurring by chance. The effects on price vary by seller. We see for the flight sellers, none of the HDPI overlap with the reference value. Differences between profiles can also be significant. Third Party 1 offers S_A prices that result in losses from \$ 4.07 to \$ 8.30 (with maximum density at \$ 6.19) compared to the control. F_B loses between \$ 0.84 and \$ 2.08 (with maximum density at \$ 1.41) compared to the control. First Party 1 offers S_A prices that result in losses from from \$ 2.40 to \$ 5.34 (with maximum density at \$ 3.68), and F_B prices that results in losses from \$ 0.46 to \$ 1.37 (with maximum density at \$ 0.83). These results suggest the expected loss is significant across a subset of profiles and sellers. For hotels, our reference profile is M_B. We see that Third Party 1015 offers the control profile prices that results in losses from \$ 2.45 to \$ 3.28 (with maximum density at \$ 2.82), while offering M_A prices that results in losses from -\$ 0.48 to -\$ 0.12 (with maximum density at -\$ 0.29); this is a better price than M_B; or a gain. Together these results show that the dollar price losses are significant. Across millions of flights sold, this results in a substantial loss to consumers.

8 DISCUSSION

Differentiation or Discrimination?: Our results show differential pricing by sellers for distinct behavioral profiles, but is the outcome discriminatory? In the US, discrimination based on protected classes is split into two categories: intentional discrimination (Section VI) and disparate impact (Section VII) [43]. To prove intentional discrimination, one must show that the seller intentionally co-varied prices based on protected attributes. To make this claim, the experimenter must be confident that they and the seller are in *consensus* regarding buyer attributes. In our framework however, the experimenter and seller can conservatively agree *only* that the behaviors are distinct. To prove disparate impact, one must show that one group received worse prices in expectation due to the seller's pricing mechanism. For example, while S_A and S_B were intended to represent senior profiles in prior work, the website lists are not associated with any real senior individuals. If we had such real data, we could make a claim about seniors. Indeed we see that these two profiles experience quite different treatment (Figures 3, 4, 5) from sellers. To make a claim about systemic disparate impact, one would need to examine how actual individuals with real browsing histories are affected as a class by a pricing mechanism. Hence, in behavior-based audits, the auditor cannot *conclude* what



(a): Third Party 1 **(b): First Party 1** **(c): First Party 2**

(d): Third Party 1015 **(e): Third Party 1016** **(f): First Party 1002**

Figure 5: *Expected Loss*: Density plots of dollar differences. The top row shows a comparison with the control profile in the flight market and the bottom row shows a comparison with $\mathrm{M_B}$ in the hotel market. The reference value is the average price difference occurring by chance between any pair of profiles — $\$\,0.44$ for flights and $\$\,0.09$ for hotels. The plots show 89% HPDI regions; non-overlapping regions imply significant differences in effects. In particular, in (a) none of the HDPI regions overlap with the reference, and many do not overlap with each other — these differences are significant (not by chance). In contrast we see in (d) that some profiles do better than the reference ($\mathrm{M_A}, \mathrm{S_A}$), while others do worse (control, $\mathrm{F_B}, \mathrm{S_B}$).

the seller *infers* from the behavior, precluding a stronger claim of intentional discrimination or disparate impact.

Comparing Markets: We observed varying levels of differential pricing in the two audited markets. Comparing the top and bottom rows in Figure 3 gives a qualitative difference between the markets. Most notably in $\mathcal{M}_F$, the control profile is on average the best performing while in $\mathcal{M}_H$, $\mathrm{M_B}$ often does well. We also see in Table 1 the coefficients for $\mathcal{M}_H$ vary less. In Figure 4, sellers that employ differential pricing to the highest degree in $\mathcal{M}_H$ still do so less than those that do the same in $\mathcal{M}_F$. In addition, the dollar impact in more extreme cases (Figure 5) is higher for flights compared to hotels. We attribute this to the general volatility of flight prices, from both short term pricing strategies [24, 55] as well as long term trends [45, 46]. These volatile conditions could potentially allow sellers to more easily employ differential pricing while remaining undetected by consumers — the method shown here can help detect this differential pricing even in high volatility scenarios, with enough data. Analysis of the results from our causal model suggests a combination of seller-dependent and seller-independent effects driving this difference between markets.

Seller Behavior: Examining Figure 4, it is clear that different sellers employ different pricing strategies to maximize profits. For example, Third Party 5 has almost an opposite strategy to Third Party 1, charging each profile more or less than the control respectively. These sellers may be making different inferences from the same behavior of our profiles, or their profit maximization could differ based on the type of business they want to attract. Before conducting these audits, we hypothesized that the control profile (with no browsing history) would see the best price, as the seller wouldn't have information with which to inform a price increase. In fact, a seller in the flight market may consider the control profile an unsophisticated or inexperienced consumer, unlikely to purchase a ticket or any add-ons (trip insurance, checked luggage, etc.), and thus not worth targeting. We see that generally, our hypothesis was correct for $\mathcal{M}_F$, with some exceptions (see Third Party 5). However we did not correctly predict the results for $\mathcal{M}_H$. One possible explanation is that price personalization may not always be in the positive direction. Individual hotels may differentiate themselves by catering to (offering lower prices to) different demographic groups. Further investigation is needed to test this hypothesis.

Designing Markets for Price Transparency: Our results demonstrate that consumers cannot be certain whether they are getting the best price when shopping in opaque markets. Sellers may take advantage of the absence of price transparency in online markets and use behavior-based differential pricing to maximize profits. Users can feel that the system is unfair [56, 63] as a result of this differential pricing.

In considering how to address this issue, we note that in most online markets, consumers are effectively "shopping alone" — the prices are highly personalized and users don't have a clear idea of what others are seeing. Price discrimination has been heavily studied, as businesses want to maximize profits by charging as much as possible. There are some instances where this can be beneficial — by utilizing auctions, sellers can elicit buyers' true valuations for goods and maximize welfare and profits. However, auctions are only feasible when supply is low and resources are constrained. In our current systems, the seller is making a guess about the consumer's willingness to pay based on some behavior which could (knowingly or unknowingly) be correlated with protected attributes. This is in contrast to auctions which try to more explicitly gauge the consumer's true willingness to pay.

One way to mitigate the impact of differential pricing is increasing market transparency; this has been shown to improve pricing in some markets [30]. In stock exchanges, the last settlement price is publicly available nearly instantaneously, and in certain markets the price is required by law to be provided. This transparency would give an inexperienced consumer some baseline of typical prices in the market. Unfortunately, a change like this may require law to take effect. Another way to make the opaque shopping experience more 'public' is by encouraging users to share information, or shop together. Users could also be invited to help shape pricing decisions; Li et al. [36] found that including consumers in the pricing design framework reduced perceptions of unfairness. Much of the behavior-based profiling we have observed in these audits relies heavily on the ad-tracking ecosystem to give sellers information about their consumers. Banning or disallowing third parties from tracking users or providing this information to sellers could potentially mitigate some of this personalized pricing – though given the centrality of advertising to the online world, this is currently an infeasible goal. More transparent structures overall will help consumers feel more confident that they're receiving a fair price.

9 LIMITATIONS AND FUTURE WORK

Other Differential Pricing Techniques: We conducted an audit on an online market with a single seller; this highlighted that sellers may implement differential pricing strategies that we do not model explicitly. We performed this audit on Macy's (macys.com) because of its variety of items and ease of querying. We found minimal differences in price across profiles. We confirmed that this is not due to lack of tracking — macys.com employs 50 distinct trackers. This, however, does not preclude Macy's from *any* differential pricing techniques. Macy's could implement personalized pricing more frequently when users are logged in, sending personalized coupons to consumers [41]. In addition, Macy's may utilize IP-based (location-based) differential pricing, something known to happen in general with e-commerce [26], but we held location constant across profiles. Finally, our experimental design may not detect differential pricing that appears due to sellers segmenting the market based on user attributes. For example, a seller may know that the men's clothing section is often visited by men. If the seller has identified this group as willing to pay a high price, then they may raise the price of all goods in the men's section compared to the women's section, regardless of who buys them. Our framework would not detect this differential pricing because we only compare equivalent goods shown to different profiles. Future work can keep these types of differential pricing strategies in mind when designing auditing frameworks.

Data: For our flight audit we only collected the first page of results; the market orders the objects by convenience. Later results show "less convenient" tickets which may include higher prices. For the same reason, we only looked at the first 5 hotel results for each of the "star" options. In our analysis, we assume stationary seller strategies over the collection period. This limitation could be mitigated by including time in the model, though this would complicate analysis.
Profiles: Constructing profiles that show an even greater difference in price is an open research question. The effect of profile size (*i.e.,* the number of websites in the profile) on the price difference is also unclear. We expect a diminishing effect of size, and we leave that for future work. Here we consider profiles defined by online behavior without being logged in to any accounts; the effect of account log-in on differential pricing is a topic for future work.

10 CONCLUSIONS

We presented a framework to audit non-transparent online markets for differential pricing by sellers using automated agents and assuming ignorance of sellers' pricing mechanisms. The notion of experimenter-seller consensus, that sellers agree on the agents' attributes set by the experimenter, is central to our framework. We conservatively implemented consensus via agent behavior, since consensus on demographics is infeasible. We audited two large online markets and showed significant causal differential pricing effects; the degree to which they occur varied with the seller. In $\mathcal{M}_F$, the control (*i.e.,* incognito) profile is almost always offered the best price — sometimes profiles are 90% more likely to see a worse price. In $\mathcal{M}_H$, $\mathrm{M_B}$ performs the best, where some profiles are 60% more likely to see a worse price. The price difference between two profiles can be as high as \$ 6.00 in $\mathcal{M}_F$ (*i.e.,* 15× the price difference by chance, \$ 0.44), and \$ 3.00 for $\mathcal{M}_H$ (*i.e.,* 33× the price difference by chance, \$ 0.09). This framework is readily extendable to other online markets and can allow for more nuanced analysis, including quantifying how sellers behave across different markets. These results demonstrate the need for more transparent pricing to inform consumers of pricing practices and mitigate unfair outcomes.

ACKNOWLEDGMENTS

We would like to thank Professors Camille Cobb, Karrie Karahalios and Gang Wang for their feedback. We also thank Adi Srinivasan, Anand Krishnan for help with hotel data collection, and Harini Arumugam and Eesha Ramkumar for qualitative analysis of profiles. Finally, thank you to the anonymous reviewers for their many useful comments and suggestions.

REFERENCES

[1] Philip Adler, Casey Falk, Sorelle A Friedler, Tionney Nix, Gabriel Rybeck, Carlos Scheidegger, Brandon Smith, and Suresh Venkatasubramanian. 2018. Auditing black-box models for indirect influence. *Knowledge and Information Systems* 54 (2018), 95–122.
[2] Pushkal Agarwal, Sagar Joglekar, Panagiotis Papadopoulos, Nishanth Sastry, and Nicolas Kourtellis. 2020. Stop Tracking Me Bro! Differential Tracking Of User Demographics On Hyper-partisan Websites. (2020). arXiv:2002.00934 [cs.CY]
[3] Julia Angwin. 2010. The web's new gold mine: Your secrets. *Wall Street Journal* 30, 07 (2010), 2010.
[4] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. 2016. Machine bias. In *Ethics of Data and Analytics*. Auerbach Publications, 254–264.
[5] Mark Armstrong. 2006. Recent developments in the economics of price discrimination. (2006).

[6] Joshua Asplund, Motahhare Eslami, Hari Sundaram, Christian Sandvig, and Karrie Karahalios. 2020. Auditing Race and Gender Discrimination in Online Housing Markets. *Proceedings of the International AAAI Conference on Web and Social Media* 14, 1 (May 2020), 24–35. https://ojs.aaai.org/index.php/ICWSM/article/view/7276

[7] Marianne Bertrand and Sendhil Mullainathan. 2004. Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American economic review* 94, 4 (2004), 991–1013.

[8] Eli Bingham, Jonathan P. Chen, Martin Jankowiak, Fritz Obermeyer, Neeraj Pradhan, Theofanis Karaletsos, Rohit Singh, Paul A. Szerlip, Paul Horsfall, and Noah D. Goodman. 2019. Pyro: Deep Universal Probabilistic Programming. *J. Mach. Learn. Res.* 20 (2019), 28:1–28:6. http://jmlr.org/papers/v20/18-403.html

[9] Robert Bloomfield and Maureen O'Hara. 1999. Market transparency: who wins and who loses? *The Review of Financial Studies* 12, 1 (1999), 5–35.

[10] Miranda Bogen and Aaron Rieke. 2018. Help wanted: An examination of hiring algorithms, equity, and bias. *Upturn, December* 7 (2018).

[11] Ralph Allan Bradley and Milton E Terry. 1952. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika* 39, 3/4 (1952), 324–345.

[12] Alicia Cackley. 2018 [Online].. Gender-Related Price Differences for Goods and Services. GAO 18-500. https://www.gao.gov/assets/gao-18-500.pdf

[13] U.S. Equal Employment Opportunity Commission. 2023. Title VII of the Civil Rights Act of 1964. retrieved from https://www.eeoc.gov/statutes/title-vii-civil-rights-act-1964.

[14] Marta R. Costa-jussà, Carlos Escolano, Christine Basta, Javier Ferrando, Roser Batlle, and Ksenia Kharitonova. 2020. Gender Bias in Multilingual Neural Machine Translation: The Architecture Matters. (2020). arXiv:2012.13176 http://arxiv.org/abs/2012.13176

[15] Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini. 2022. Who Audits the Auditors? Recommendations from a Field Scan of the Algorithmic Auditing Ecosystem. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1571–1583. https://doi.org/10.1145/3531146.3533213

[16] Ruomeng Cui, Jun Li, and Dennis J Zhang. 2020. Reducing discrimination with reviews in the sharing economy: Evidence from field experiments on Airbnb. *Management Science* 66, 3 (2020), 1071–1094.

[17] Amit Datta, Michael Carl Tschantz, and Anupam Datta. 2015. Automated Experiments on Ad Privacy Settings. *Proceedings on Privacy Enhancing Technologies* 2015, 1 (2015), 92–112. https://doi.org/doi:10.1515/popets-2015-0007

[18] Steve Deane. 2022. Over 60 Online Travel Booking Statistics (2022). https://www.stratosjets.com/blog/online-travel-statistics/

[19] Jean-Pierre Dubé and Sanjog Misra. 2017. *Personalized Pricing and Consumer Welfare.* Working Paper 23775. National Bureau of Economic Research. https://doi.org/10.3386/w23775

[20] Steven Englehardt, Dillon Reisman, Christian Eubank, Peter Zimmerman, Jonathan Mayer, Arvind Narayanan, and Edward W Felten. 2015. Cookies that give you away: The surveillance implications of web tracking. In *Proceedings of the 24th International Conference on World Wide Web*. 289–299.

[21] Henry S Farber, Dan Silverman, and Till Von Wachter. 2016. Determinants of callbacks to job applications: An audit study. *American Economic Review* 106, 5 (2016), 314–318.

[22] Yunhe Feng and Chirag Shah. 2022. Has CEO Gender Bias Really Been Fixed? Adversarial Attacking and Improving Gender Fairness in Image Search. In *Proceedings of the AAAI conference on artificial intelligence (AAAI)*.

[23] Michael Geden and Joshua Andrews. 2021. Fair and Interpretable Algorithmic Hiring Using Evolutionary Many-Objective Optimization. *35th AAAI Conference on Artificial Intelligence, AAAI 2021* 17A (2021), 14795–14803.

[24] David Gillen and Benny Mantin. 2009. Price volatility in the airline markets. *Transportation Research Part E: Logistics and Transportation Review* 45, 5 (2009), 693–709. https://doi.org/10.1016/j.tre.2009.04.014

[25] Aniko Hannak, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson. 2014. Measuring Price Discrimination and Steering on E-Commerce Web Sites. In *Proceedings of the 2014 Conference on Internet Measurement Conference* (Vancouver, BC, Canada) *(IMC '14)*. Association for Computing Machinery, New York, NY, USA, 305–318. https://doi.org/10.1145/2663716.2663744

[26] Christoph Michael Hindermann. 2018. Price discrimination in online retail. (2018).

[27] Matthew D Hoffman, David M Blei, Chong Wang, and John Paisley. 2013. Stochastic variational inference. *Journal of Machine Learning Research* (2013).

[28] Thomas Hupperich, Dennis Tatang, Nicolai Wilkop, and Thorsten Holz. 2018. An empirical study on online price differentiation. In *Proceedings of the eighth ACM conference on data and application security and privacy*. 76–83.

[29] Basileal Imana, Aleksandra Korolova, and John Heidemann. 2021. Auditing for discrimination in algorithms delivering job ads. In *Proceedings of the Web Conference 2021*. 3767–3778.

[30] Elena Ionascu, Paloma Taltavull de La Paz, and Marilena Mironiuc. 2021. The Relationship between Housing Prices and Market Transparency. Evidence from the Metropolitan European Markets. *Housing, Theory and Society* 38, 1 (2021), 42–71.

[31] Costas Iordanou, Claudio Soriente, Michael Sirivianos, and Nikolaos Laoutaris. 2017. Who is Fiddling with Prices? Building and Deploying a Watchdog Service for E-Commerce. In *Proceedings of the Conference of the ACM Special Interest Group on Data Communication* (Los Angeles, CA, USA) *(SIGCOMM '17)*. Association for Computing Machinery, New York, NY, USA, 376–389. https://doi.org/10.1145/3098822.3098850

[32] Rashidul Islam, Kamrun Naher Keya, Ziqian Zeng, Shimei Pan, and James Foulds. 2021. Debiasing career recommendations with neural fair collaborative filtering. In *Proceedings of the Web Conference 2021*. 3779–3790.

[33] Lina Khan. 2017. Amazon's antitrust paradox. *Yale law journal* 126 (2017).

[34] Juhi Kulshrestha, Marcos Oliveira, Orkut Karaçalık, Denis Bonnay, and Claudia Wagner. 2021. Web routineness and limits of predictability: Investigating demographic and behavioral differences using web tracking data. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 15. 327–338.

[35] Karen Levy and Solon Barocas. 2017. Designing against discrimination in online markets. *Berkeley Technology Law Journal* 32, 3 (2017), 1183–1238.

[36] Wenjing Li, David M Hardesty, and Adam W Craig. 2018. The impact of dynamic bundling on price fairness perceptions. *Journal of Retailing and Consumer Services* 40 (2018), 204–212.

[37] Sean Matthews, John Hudzina, and Dawn Sepehr. 2022. Gender and Racial Stereotype Detection in Legal Opinion Word Embeddings. *arXiv preprint arXiv:2203.13369* (2022).

[38] Jakub Mikians, László Gyarmati, Vijay Erramilli, and Nikolaos Laoutaris. 2012. Detecting price and search discrimination on the Internet. *Proceedings of the 11th ACM Workshop on Hot Topics in Networks, HotNets-11* July (2012), 79–84. https://doi.org/10.1145/2390231.2390245

[39] Rafi Mohammed. 2017. How retailers use personalized prices to test what you're willing to pay. *Harvard Business Review* 20 (2017).

[40] Faidra Georgia Monachou and Itai Ashlagi. 2019. Discrimination in online markets: Effects of social bias on learning from reviews and policy design. *Advances in Neural Information Processing Systems* 32 (2019).

[41] Chakravarthi Narasimhan. 1984. A price discrimination theory of coupons. *Marketing Science* 3, 2 (1984), 128–147.

[42] U.S. Department of Housing and Urban Development. 2023. History of Fair Housing. retrieved from https://www.hud.gov/program_offices/fair_housing_equal_opp/aboutfheo/history.

[43] U.S Department of Justice. 2021. Title VI Legal Manual. retrieved from https://www.justice.gov/crt/book/file/1364106/download.

[44] U.S. Department of Justice. 2023. FAIR HOUSING TESTING PROGRAM. retrieved from https://www.justice.gov/crt/fair-housing-testing-program-1.

[45] U.S. Bureau of Labor Statistics. 2023. Consumer Price Index for All Urban Consumers: Airline Fares in U.S. City Average [CUSR0000SETG01]. retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/CUSR0000SETG01.

[46] U.S. Bureau of Labor Statistics. 2023. Consumer Price Index for All Urban Consumers: Lodging Away from Home in U.S. City Average [CUSR0000SEHB]. retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org/series/CUSR0000SEHB.

[47] Sun Jung Oh and John Yinger. 2015. What have we learned from paired testing in housing markets? *Cityscape* 17, 3 (2015), 15–60.

[48] Cathy O'neil. 2016. *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway books.

[49] Panagiotis Papadopoulos, Nicolas Kourtellis, and Evangelos Markatos. 2019. Cookie Synchronization: Everything You Always Wanted to Know But Were Afraid to Ask. In *The World Wide Web Conference* (San Francisco, CA, USA) *(WWW '19)*. ACM, New York, NY, USA, 1432–1442. https://doi.org/10.1145/3308558.3313542

[50] Saerom Park, Seongmin Kim, and Yeon-sup Lim. 2022. Fairness Audit of Machine Learning Models with Confidential Computing. In *Proceedings of the ACM Web Conference 2022*. 3488–3499.

[51] Judea Pearl. 2009. *Causality*. Cambridge university press.

[52] Andi Peng, Besmira Nushi, Emre Kiciman, Kori Inkpen, and Ece Kamar. 2022. Investigations of Performance and Bias in Human-AI Teamwork in Hiring. *arXiv preprint arXiv:2202.11812* (2022).

[53] Du Phan, Neeraj Pradhan, and Martin Jankowiak. 2019. Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro. *arXiv preprint arXiv:1912.11554* (2019).

[54] Lincoln Quillian, John J Lee, and Brandon Honoré. 2020. Racial discrimination in the US housing and mortgage lending markets: a quantitative review of trends, 1976–2016. *Race and Social Problems* 12 (2020), 13–28.

[55] Retuers. 2013. Airlines versus hotels in the battle for the best pricing strategy: Travel industry news; conferences - reuters events. https://www.reutersevents.com/travel/revenue-and-data-management/airlines-versus-hotels-battle-best-pricing-strategy

[56] Timothy J Richards, Jura Liaukonyte, and Nadia A Streletskaya. 2016. Personalized pricing and price fairness. *International Journal of Industrial Organization* 44 (2016), 138–153.

[57] Antonio Scalia. 1999. Does market transparency matter? A case study. *A Case Study (October 10, 1999). Bank of Italy Temi di Discussione (Working Paper) No* 359 (1999).

[58] Janice C Sipior, Burke T Ward, and Ruben A Mendoza. 2011. Online privacy concerns associated with cookies, flash cookies, and web beacons. *Journal of Internet Commerce* 10, 1 (2011), 1–16.

[59] Lars A Stole. 2007. Price discrimination and competition. *Handbook of industrial organization* 3 (2007), 2221–2299.

[60] Sotiris Tsioutsiouliklis, Evaggelia Pitoura, Panayiotis Tsaparas, Ilias Kleftakis, and Nikos Mamoulis. 2021. Fairness-aware pagerank. In *Proceedings of the Web Conference 2021*. 3815–3826.

[61] Hal R. Varian. 1989. Chapter 10 Price discrimination. Handbook of Industrial Organization, Vol. 1. Elsevier, 597–654. https://doi.org/10.1016/S1573-448X(89)01013-7

[62] Thomas Vissers, Nick Nikiforakis, Nataliia Bielova, and Wouter Joosen. 2014. Crying wolf? on the price discrimination of online airline tickets. In *7th Workshop on Hot Topics in Privacy Enhancing Technologies (HotPETs 2014)*.

[63] Fei L Weisstein, Kent B Monroe, and Monika Kukar-Kinney. 2013. Effects of price framing on consumers' perceptions of online dynamic pricing practices. *Journal of the Academy of Marketing Science* 41 (2013), 501–514.

[64] Ronald E Wienk. 1979. *Measuring racial discrimination in American housing markets: The housing market practices survey*. Vol. 444. Division of Evaluation, US Department of Housing and Urban Development.

[65] Christo Wilson, Avijit Ghosh, Shan Jiang, Alan Mislove, Lewis Baker, Janelle Szary, Kelly Trindel, and Frida Polli. 2021. Building and Auditing Fair Algorithms: A Case Study in Candidate Screening. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (Virtual Event, Canada) *(FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 666–677. https://doi.org/10.1145/3442188.3445928

[66] Sewall Wright. 1920. The Relative Importance of Heredity and Environment in Determining the Piebald Pattern of Guinea-Pigs. *Proceedings of the National Academy of Sciences* 6, 6 (1920), 320–332. https://doi.org/10.1073/pnas.6.6.320 arXiv:https://www.pnas.org/doi/pdf/10.1073/pnas.6.6.320

[67] Renzhe Xu, Xingxuan Zhang, Peng Cui, Bo Li, Zheyan Shen, and Jiazheng Xu. 2022. Regulatory Instruments for Fair Personalized Pricing. In *Proceedings of the ACM Web Conference 2022*. 4–15.

[68] John Yinger. 1986. Measuring racial discrimination with fair housing audits: Caught in the act. *The American Economic Review* (1986), 881–893.

[69] Meg Young, Michael Katell, and P.M. Krafft. 2022. Confronting Power and Corporate Capture at the FAccT Conference. In *2022 ACM Conference on Fairness, Accountability, and Transparency* (Seoul, Republic of Korea) *(FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 1375–1386. https://doi.org/10.1145/3531146.3533194

[70] Frederik Zuiderveen Borgesius and Joost Poort. 2017. Online price discrimination and EU data privacy law. *Journal of consumer policy* 40, 3 (2017), 347–366.

A FULL RESULTS FOR DIFFERENTIAL TREATMENT (RQ1)

In the main body of the paper (Section 7.1) we show differential treatment across profiles for select sellers. We chose to only show select sellers to save space, but displayed the 89% HDI regions for each profile. Here we give the mean value $p_{i,c|k}$ for all profiles $a_i \in A$ and all sellers $k \in S$ for $\mathcal{M}_F$. This is the probability of profile a_i seeing a worse price than the control profile C (Table 2a). We are unable to show the full table for $\mathcal{M}_H$ in this appendix because of the large number of sellers we observed in this market, however we do present an extended table (Table 2b) compared to the one in the main paper. We present the mean value $p_{i,j|k}$ for all profiles $a_i \in A$ where $a_j = \mathrm{M_B}$. In both markets, sellers' pricing behavior can vary dramatically — in $\mathcal{M}_F$, Third Party 5 shows very low probabilities across profiles (ranging from 1% – 11%) while others like Third Party 18 show a larger range (18% – 80%). In $\mathcal{M}_H$ the range of behaviors still exist but the effect is more muted - Third Party 1010 ranges from (6% – 29%) while most others are similar to the base probability (*e.g.,* First party 1356, First Party 1361). Though the original authors of our profiles assign demographic identities to these profiles, equivalent identities from each profile set do not experience similar effects. In particular, in both $\mathcal{M}_F$ and $\mathcal{M}_H$, $\mathrm{F_A}$ often experiences a very high probability of seeing a worse price while for the same sellers, $\mathrm{F_B}$ experiences a lower probability of seeing a worse price.

In Section 7.1 we also asked how much of the differential treatment between profile a_i and the control C is due to the seller, profile interaction *alone*. We answer this question by examining the interaction effect $\delta_{k,i}$ for all sellers k and profiles a_i. We showed a truncated version of the figure in the main body, and show the full version for $\mathcal{M}_F$ in Figure 7a and expanded version for $\mathcal{M}_H$ in Figure 7b.

Keep in mind that this result must be combined with the seller-independent effect to understand the overall effect of differential treatment. However, even with these general trends for $\delta_{k,i}$, the seller specific trends can vary quite a bit — Third Party 7 and Third Party 14 have almost opposite strategies (that is, a given profile will experience a positive coefficient for one seller and a negative for another).

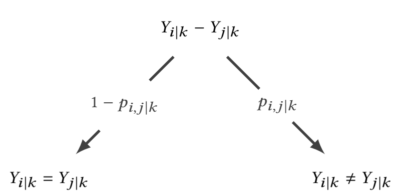


Figure 6: The input to this model is a dataset where each row is a price comparison (in dollars) between profile a_i and a_j, conditioned on a good g and seller k, $Y_{i|k} - Y_{j|k}$. With probability $p_{i,j|k}$, profiles a_i and a_j see different prices. With probability $1 - p_{i,j|k}$, they see the same price. $p_{i,j|k}$ is the output from our original model (see Equation (2)).

B PRICE DIFFERENCE MODEL AND RESULTS (RQ2)

First we depict the mixture model we described in the main body, which has the aggregate effect of a zero-inflated Gaussian.

We model the outcome $Y_{i,j|k}$ with a normal distribution with parameters $\mu_{i,j|k}$ and $\sigma_{i,j|k}$ (Equation (9)). $\mu_{i,j|k}$ gives the mean price difference (in dollars) between profiles a_i, a_j conditioned on the same seller k (Equation (10)). $\mu_{i,j|k}$ is also conditioned on the good g; we've avoided conditioning on g in the equations for clarity.

On the right-hand side of Equation (10), $\bar{p}$ is the base dollar difference between tickets seen by profiles a_i and a_j. We can model the effect of the profile a_i through the unobserved confound U_1 with a parameter β_i. Thus these parameters affect the dollar difference through the difference in their direct effects, $(\beta_i - \beta_j)$. We capture the seller-profile interaction effect through the parameter $\delta_{k,i}$. Thus, we introduce $\delta_{k,i} - \delta_{k,j}$, the difference of the interaction effects to understand their effect on the price difference.

Equation (12) shows our multi-level modeling for the direct effects (β). We draw the β_i from a common prior $\mathrm{N}(\bar{\beta}, \beta_\sigma)$ allowing for pooling, thus regularizing our estimates for β_i. We use weakly informative priors for both $\bar{\beta}$ and β_σ, as well as $\delta_{i,k}$ (seller profile interaction shown in Equation (15)). Notice that we set the mean of the prior for both $\bar{\beta}$ and $\delta_{i,k}$ to be 0, because a priori, we do not expect either a direct effect of the profile or an effect of the interaction between seller and profile.

Finally, Equation (16) shows the prior for $\bar{p}$, which gives the base dollar difference. We set the mean value to \$5 because we know that the average price difference between profiles (when it exists) is approximately \$5. We set the standard deviation to \$30 because we also know that the difference can vary substantially.

$$Y_{i,j|k} \sim \mathrm{N}(\mu_{i,j|k}, \sigma_{i,j|k}) \tag{9}$$
$$\mu_{i,j|k} = \bar{p} + (\beta_i - \beta_j) + (\delta_{k,i} - \delta_{k,j}) \tag{10}$$
$$\sigma_{i,j|k} \sim \mathrm{N}(0.0, 1.0) \tag{11}$$
$$\beta_i \sim \mathrm{N}(\bar{\beta}, \beta_\sigma) \tag{12}$$
$$\bar{\beta} \sim \mathrm{N}(0.0, 1.0) \tag{13}$$
$$\beta_\sigma \sim \mathrm{Exp}(1.0) \tag{14}$$
$$\delta_{k,i} \sim \mathrm{N}(0.0, 1.0) \tag{15}$$
$$\bar{p} \sim \mathrm{N}(5.0, 30.0) \tag{16}$$

B.1 Results

In the main body of the paper, we show the expected loss of each profile compared to the best performing profile for select sellers. We obtain these values by multiplying the expected loss conditioned on the existence of a price difference by the probability of a price difference, as depicted in Figure 6. Here we show the plots for the expected loss conditioned on the existence of a price difference (Figure 8), which is the output of the model above. We show the same selected sellers for $\mathcal{M}_F$ and $\mathcal{M}_H$ as presented in the paper. We note that the range of the conditional loss can be quite high. In Figure 8c for First Party 2, the loss compared to the control ranges from \$ 5.50 to \$ 16.00, while for Third Party 1015 in Figure 8e the loss compared to $\mathrm{M_B}$ ranges from -\$ 2.00 to \$ 8.00.

We also show the mean price difference over all goods g, $\mathbb{E}(\rho(g_i) - \rho(g_j))$, where $g_i \equiv g_j$ for all sellers $k \in S$ and profiles $a_i \in A$ in $\mathcal{M}_F$ (Table 3a). We show a condensed version of $\mathcal{M}_H$ for space (Table 3b). This is the mean loss (in dollars) a_i incurs when compared to the control profile C and M_B respectively, and are calculated as described in Figure 6. The range of impacts are noteworthy. In $\mathcal{M}_F$, Third Party 5 has fairly low and consistent dollar losses while First Party 7 has quite disparate losses (ranging from \$ 1.17 to \$ 8.76). Profiles meant to represent equivalent identities from each profile set do not experience similar effects. In particular, M_A often experiences a higher expected loss per seller, relative to M_B. Also notably, in $\mathcal{M}_H$ we generally see lower impacts across sellers and profiles. For example, the loss caused by Third Party 1015 varies from -\$ 0.30 to \$ 2.86 while that caused by First Party 1230 ranges from \$ 0.04 to \$ 0.11.

C DATA DETAILS

In this section we give additional details about the data we collected from kayak.com.

Each record we collected from $\mathcal{M}_F$ contains the following information: the departure and arrival information (time, airport, carrier, and the number of stops), the seller of the ticket, and the price. Each record we collected from $\mathcal{M}_H$ contains the following information: the departure and arrival dates, the city, hotel, seller, room configuration, any amenities, and the price.

Notably, in both cases, the seller of the ticket may not be the flight carrier or hotel owner, as aggregators can sell seats for other airlines or rooms in other hotels. We distinguish between first party sellers (those who sell tickets on their own branded planes or rooms in their own branded hotels) and third party sellers (aggregators or other ticketing services that do not operate any actual flights or own any hotels). To anonymize our data, we do not list specific cities, routes or sellers. Instead, we characterize sellers as first or third party, cities as large, medium, or small, and airports as large hub, medium hub, small hub, or non-hub. Thus the route types are every combination of airport types, except non-hub - non-hub, because we did not query such routes. See Figure 9 for descriptive statistics of both flight data (9a) and hotel data (9b).

Now we give a few details regarding the implementation of the data collection algorithm — for every query $q \in Q$ (Algorithm 1, line 1), we train agents $i \in \mathcal{N}$ (lines 1–1) using Selenium (with Firefox) by visiting websites $w_i \in W$, thus creating each agent's profile a_i. Each agent is trained in parallel on a machine with enough cores such that each process can execute on individual cores. No shared memory is used. Once all agents are trained, all processes are started simultaneously. The queries themselves may not execute exactly at the same time (due to process scheduling, etc.), but the ultimate order is randomized since any delay-causing factors are independent of the market conditions and agent identity, and thus do not affect our results. To ensure consistency (location and time each day) in queries, we run each batch of queries on a university Linux cluster.

Table 2: *Differential Treatment:* Mean value of $p_{i,c|k}$ for all sellers k and profiles a_i when compared to the reference profile.

(a) *Differential Treatment* $\mathcal{M}_F$: Reference profile is C **The variation across sellers is noteworthy – Third Party 5 shows very low probabilities across profiles (ranging from 1% – 11%) while others like Third Party 18 show quite high variation range (**18% – 80%**).**

Seller	F_A	F_B	M_A	M_B	S_A	S_B	Y_A	Y_B
First Party 1	0.34	0.15	0.48	0.26	0.58	0.18	0.25	0.34
First Party 2	0.15	0.10	0.26	0.14	0.17	0.08	0.09	0.16
First Party 3	0.45	0.16	0.66	0.31	0.68	0.18	0.68	0.49
First Party 4	0.26	0.11	0.21	0.15	0.47	0.21	0.19	0.28
First Party 5	0.19	0.11	0.37	0.16	0.22	0.09	0.18	0.21
First Party 6	0.44	0.15	0.61	0.30	0.73	0.20	0.32	0.47
First Party 7	0.77	0.25	0.90	0.51	0.91	0.33	0.47	0.80
First Party 8	0.52	0.19	0.71	0.36	0.74	0.22	0.87	0.56
First Party 9	0.37	0.14	0.51	0.26	0.66	0.19	0.28	0.40
Third Party 1	0.87	0.26	0.94	0.61	0.97	0.44	0.38	0.88
Third Party 2	0.56	0.20	0.75	0.39	0.77	0.23	0.79	0.60
Third Party 3	0.43	0.15	0.62	0.29	0.62	0.17	0.40	0.47
Third Party 4	0.56	0.18	0.75	0.41	0.79	0.25	0.60	0.61
Third Party 5	0.04	0.11	0.05	0.09	0.01	0.04	0.04	0.05
Third Party 6	0.11	0.11	0.17	0.16	0.12	0.07	0.02	0.12
Third Party 7	0.06	0.09	0.15	0.10	0.06	0.05	0.09	0.08
Third Party 8	0.63	0.22	0.81	0.40	0.82	0.25	0.14	0.67
Third Party 9	0.37	0.14	0.51	0.25	0.63	0.18	0.27	0.38
Third Party 10	0.58	0.19	0.75	0.37	0.77	0.20	0.32	0.60
Third Party 11	0.10	0.10	0.18	0.15	0.11	0.07	0.02	0.11
Third Party 12	0.62	0.21	0.80	0.41	0.85	0.25	0.52	0.65
Third Party 13	0.84	0.26	0.93	0.60	0.96	0.43	0.59	0.86
Third Party 14	0.67	0.22	0.83	0.46	0.85	0.29	0.45	0.71
Third Party 15	0.45	0.16	0.61	0.31	0.74	0.21	0.33	0.48
Third Party 16	0.49	0.17	0.69	0.34	0.70	0.20	0.40	0.53
Third Party 17	0.26	0.11	0.21	0.14	0.47	0.22	0.18	0.28
Third Party 18	0.56	0.18	0.75	0.38	0.80	0.23	0.38	0.60
Third Party 19	0.22	0.11	0.37	0.19	0.33	0.11	0.35	0.25
Third Party 20	0.38	0.14	0.52	0.26	0.65	0.18	0.27	0.41
Third Party 21	0.25	0.13	0.41	0.21	0.34	0.12	0.16	0.27
Third Party 22	0.84	0.26	0.93	0.59	0.96	0.42	0.51	0.86
Third Party 23	0.39	0.15	0.58	0.28	0.57	0.16	0.43	0.42
Third Party 24	0.60	0.19	0.79	0.39	0.83	0.24	0.50	0.64
Third Party 25	0.54	0.19	0.73	0.36	0.77	0.22	0.38	0.58

(b) *Differential Treatment* $\mathcal{M}_H$: Reference profile compared with M_B**. The variation across sellers is noteworthy — First Party 1230 shows very low probabilities across profiles (ranging from 1% – 3%) while others like Third Party 1010 show quite high variation range (**6% – 29%**).**

Seller	F_A	F_B	M_A	S_A	S_B	Y_A	Y_B	C
First Party 1002	0.07	0.05	0.05	0.05	0.04	0.03	0.09	0.06
First Party 1069	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1092	0.01	0.01	0.01	0.02	0.01	0.02	0.06	0.02
First Party 1229	0.03	0.05	0.04	0.03	0.05	0.03	0.04	0.04
First Party 1230	0.01	0.02	0.02	0.01	0.03	0.01	0.03	0.01
First Party 1255	0.01	0.03	0.03	0.03	0.07	0.02	0.1	0.07
First Party 1259	0.05	0.1	0.13	0.06	0.1	0.04	0.16	0.25
First Party 1278	0.05	0.16	0.08	0.05	0.08	0.03	0.14	0.15
First Party 1291	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1298	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.06
First Party 1301	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1310	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1318	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1338	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1356	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1361	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1362	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1364	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1380	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1391	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1392	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1396	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
First Party 1397	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.06
First Party 1408	0.05	0.06	0.04	0.06	0.05	0.03	0.04	0.05
Third Party 1001	0.02	0.06	0.03	0.02	0.05	0.02	0.08	0.10
Third Party 1005	0.03	0.07	0.04	0.03	0.05	0.03	0.09	0.14
Third Party 1007	0.03	0.07	0.04	0.03	0.06	0.03	0.09	0.15
Third Party 1009	0.03	0.07	0.04	0.03	0.06	0.03	0.09	0.15
Third Party 1010	0.06	0.13	0.07	0.05	0.09	0.05	0.14	0.29
Third Party 1011	0.03	0.06	0.04	0.02	0.05	0.02	0.08	0.15
Third Party 1015	0.15	0.33	0.21	0.13	0.25	0.13	0.36	0.59
Third Party 1016	0.03	0.08	0.05	0.03	0.06	0.03	0.11	0.28
Third Party 1017	0.03	0.08	0.05	0.03	0.06	0.03	0.1	0.16
Third Party 1020	0.03	0.08	0.05	0.03	0.06	0.03	0.1	0.15

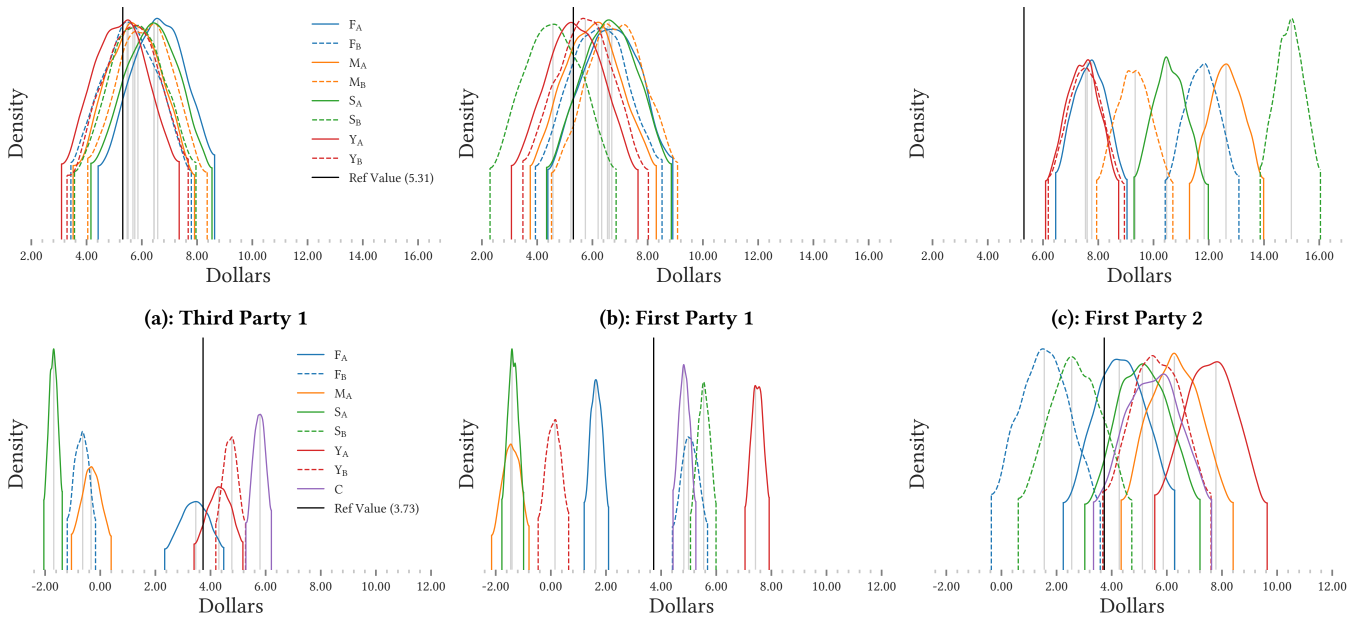
	C	F_A	F_B	M_A	M_B	S_A	S_B	Y_A	Y_B
Third Party 17	0.00	0.01	-0.01	0.02	-0.00	0.01	-0.01	-0.00	0.02
First Party 4	-0.00	-0.00	0.01	0.01	0.04	0.01	-0.06	-0.00	0.01
First Party 1	-0.42	-0.04	-0.03	0.88	0.32	0.07	-0.67	-0.00	-0.07
Third Party 9	-0.49	0.03	-0.19	0.90	0.23	0.22	-0.71	-0.00	0.01
First Party 9	-0.52	0.00	-0.22	0.90	0.23	0.28	-0.73	0.01	0.06
Third Party 20	-0.52	0.04	-0.25	0.90	0.25	0.27	-0.77	-0.00	0.11
Third Party 21	0.09	0.04	0.31	1.08	0.56	-0.45	-0.65	-0.08	0.08
Third Party 19	-0.29	-0.47	-0.28	0.54	0.07	-0.86	-1.09	0.59	-0.39
First Party 6	-0.74	0.08	-0.34	1.07	0.23	0.42	-0.84	-0.00	0.13
Third Party 15	-0.78	0.10	-0.32	1.05	0.21	0.43	-0.84	0.00	0.16
First Party 2	0.27	-0.41	0.24	0.56	0.23	-1.19	-0.89	-0.53	-0.37
First Party 5	0.26	-0.15	0.27	1.08	0.41	-0.85	-0.81	0.25	-0.09
Third Party 23	-0.83	-0.22	-0.46	0.84	0.06	-0.41	-1.19	0.42	-0.17
Third Party 3	-0.55	0.22	-0.16	1.30	0.35	0.08	-0.88	0.58	0.31
Third Party 16	-0.74	0.27	-0.17	1.44	0.39	0.26	-0.86	0.39	0.39
Third Party 25	-0.30	0.92	0.35	2.07	0.95	1.05	-0.27	0.73	1.01
Third Party 18	-0.94	0.38	-0.30	1.53	0.37	0.60	-0.85	0.06	0.47
Third Party 10	-0.87	0.51	-0.18	1.60	0.40	0.49	-0.96	-0.11	0.54
Third Party 4	-1.46	-0.16	-0.86	0.97	-0.01	-0.01	-1.27	0.46	0.00
First Party 3	-0.92	-0.04	-0.41	1.11	0.11	-0.00	-1.14	1.37	0.05
Third Party 24	-1.59	-0.11	-0.93	1.08	-0.21	0.15	-1.44	-0.08	-0.00
Third Party 12	-1.53	0.02	-0.72	1.22	-0.08	0.38	-1.35	0.05	0.10
Third Party 7	1.22	-0.39	1.03	0.80	0.88	-1.44	-0.48	0.37	-0.23
Third Party 14	-1.11	0.64	-0.22	1.87	0.55	0.76	-0.71	0.21	0.78
Third Party 6	0.40	-0.60	0.44	0.20	0.56	-1.41	-0.94	-2.14	-0.60
Third Party 11	-0.04	-1.18	-0.05	-0.23	0.02	-2.03	-1.40	-2.53	-1.16
Third Party 2	-1.50	-0.18	-0.77	0.95	-0.14	-0.15	-1.41	1.38	-0.08
Third Party 8	-1.31	0.30	-0.43	1.49	0.12	0.39	-1.11	-1.59	0.40
First Party 8	-0.68	0.45	-0.02	1.58	0.56	0.53	-0.68	2.76	0.57
First Party 7	-1.81	0.46	-0.79	1.74	0.02	0.71	-1.26	-0.43	0.60
Third Party 13	-1.59	1.12	-0.50	2.36	0.63	1.85	-0.60	0.29	1.26
Third Party 5	1.19	-0.95	1.19	-0.33	0.69	-2.92	-0.68	-0.42	-0.82
Third Party 22	-1.96	0.76	-0.86	2.05	0.23	1.42	-1.00	-0.42	0.90
Third Party 1	-1.59	1.39	-0.48	2.55	0.68	1.89	-0.55	-0.54	1.45

(a): ***Seller effects*** **$\mathcal{M}_F$: While there are some profiles that have a tendency to see the same direction of effect across many sellers, (Control, M_A, S_B) others have a bit more variation.**

	C	F_A	F_B	M_A	M_B	S_A	S_B	Y_A	Y_B
First Party 1356	-0.00	0.00	-0.01	0.00	-0.00	0.00	-0.01	-0.00	0.00
First Party 1301	0.00	0.01	0.00	0.01	0.01	0.00	0.01	0.01	-0.00
First Party 1338	0.00	-0.00	-0.01	-0.01	0.00	-0.00	0.01	-0.01	-0.00
First Party 1396	0.00	-0.00	-0.00	-0.00	-0.00	0.01	0.01	-0.00	0.00
First Party 1310	-0.01	-0.00	0.00	-0.01	-0.00	0.01	-0.00	0.01	0.00
First Party 1362	-0.00	-0.00	-0.01	-0.00	-0.00	0.00	-0.00	-0.01	-0.01
First Party 1391	0.00	-0.00	-0.00	-0.01	0.00	0.01	0.00	-0.01	-0.00
First Party 1318	0.00	-0.00	-0.01	-0.00	-0.01	0.00	-0.01	0.00	0.00
First Party 1069	-0.00	-0.01	0.01	-0.01	-0.00	-0.00	-0.00	0.00	0.00
First Party 1380	0.01	-0.00	-0.00	-0.00	0.00	0.00	0.01	0.00	0.01
First Party 1364	-0.01	-0.01	0.00	-0.00	-0.01	0.01	-0.00	0.01	-0.00
First Party 1397	0.00	0.00	-0.01	0.00	-0.01	0.01	0.00	0.01	0.00
First Party 1291	-0.00	-0.00	-0.00	-0.00	-0.01	-0.00	-0.00	-0.02	0.01
First Party 1408	-0.01	-0.01	0.00	0.00	0.01	-0.00	0.01	0.00	0.01
First Party 1298	-0.00	-0.00	-0.01	-0.00	-0.01	-0.00	0.00	0.01	0.00
First Party 1361	0.00	0.01	0.01	-0.01	-0.00	0.00	0.01	-0.00	-0.00
First Party 1392	0.01	-0.01	-0.00	0.00	-0.00	-0.01	-0.00	-0.02	-0.00
Third Party 1002	0.81	-0.59	0.27	-0.01	-0.13	-1.02	0.14	-0.38	0.80
First Party 1278	0.61	-0.49	0.61	0.15	-0.57	-0.84	0.09	-0.56	0.80
Third Party 1005	0.87	-0.71	0.06	-0.15	-0.22	-1.11	-0.00	-0.43	0.65
Third Party 1020	1.68	0.14	0.86	0.67	0.48	-0.28	0.82	0.35	1.50
First Party 1228	0.32	-1.28	0.43	-0.21	0.56	-0.66	0.43	-0.04	0.47
First Party 1230	-0.87	-0.32	-0.36	-0.06	0.91	-0.74	0.53	0.06	0.70
Third Party 1007	0.65	-0.95	-0.21	-0.42	-0.54	-1.37	-0.26	-0.74	0.42
Third Party 1017	0.68	-0.92	-0.19	-0.39	-0.55	-1.34	-0.23	-0.70	0.44
Third Party 1009	0.60	-1.00	-0.27	-0.47	-0.60	-1.43	-0.32	-0.79	0.36
Third Party 1001	1.02	-0.47	0.49	0.20	0.34	-0.80	0.39	-0.25	1.07
Third Party 1011	0.61	-1.13	-0.41	-0.58	-0.54	-1.53	-0.42	-0.92	0.25
First Party 1259	1.08	-0.55	-0.12	0.58	-0.72	-0.70	0.16	-0.54	0.80
Third Party 1010	1.38	-0.33	0.29	0.02	-0.63	-0.93	0.10	-0.18	0.76
Third Party 1016	1.89	-0.38	0.36	0.23	-0.06	-0.72	0.34	-0.02	1.04
First Party 1092	0.08	-0.89	-0.83	-0.57	0.91	-0.14	-0.49	0.56	1.33
First Party 1255	0.59	-1.58	-0.46	-0.16	0.23	-0.66	0.75	-0.03	1.23
Third Party 1015	1.57	-0.39	0.46	0.16	-1.72	-0.86	0.29	-0.19	0.92

(b): ***Seller effects*** **$\mathcal{M}_H$: We see that sellers can have very differentiated strategies — First Party 1230 and Third Party 1015 treat the Control, F_A, M_B, Y_B in opposing directions.**

Figure 7: ***Seller effects*****: We show the interaction variable $\delta_{k,i}$ for a further subset of sellers, ordered from lowest to highest variance within the row. Note that the strategy across sellers (what each sellers shows a profile) is variable.**



(a): Third Party 1 **(b): First Party 1** **(c): First Party 2**

(d): Third Party 1016 **(e): Third Party 1015** **(f): First Party 1002**

Figure 8: *Expected Loss Conditioned on Existence of Price Differences*: Density plots of price differences in dollars conditioned on price differences. The top three in $\mathcal{M}_F$ are compared with the control profile while the bottom three in $\mathcal{M}_H$ are compared with $\mathrm{M_B}$. The reference value is the average price difference \$ 5.31 for M_F and \$ 3.73 for M_H. The plots show 89% HPDI regions; non-overlapping regions imply significant differences in effects. For example, First Party 2 (c) in $\mathcal{M}_F$ shows M_B and F_A have non-significant overlap among others, while for Third Party 1015 (e) in $\mathcal{M}_H$, $\mathrm{S_B}$ and $\mathrm{Y_B}$ are non-overlapping.

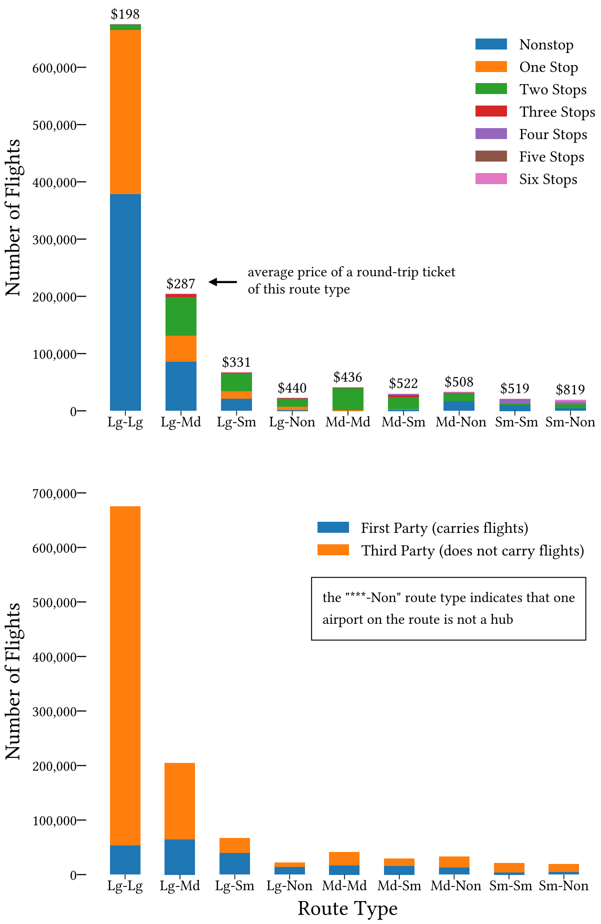
Table 3: *Expected Dollar Loss:* Mean loss in dollars for sellers and profiles compared to the reference profile.

(a) *Expected Dollar Loss* $\mathcal{M}_F$: Mean loss in dollars for all sellers and profiles compared with C. The range of impacts are noteworthy. Profiles who buy from Third Party 5 have fairly low and consistent dollar losses while those who buy from First Party 7 have quite disparate losses — ranging from $ 1.17 to $ 8.76.

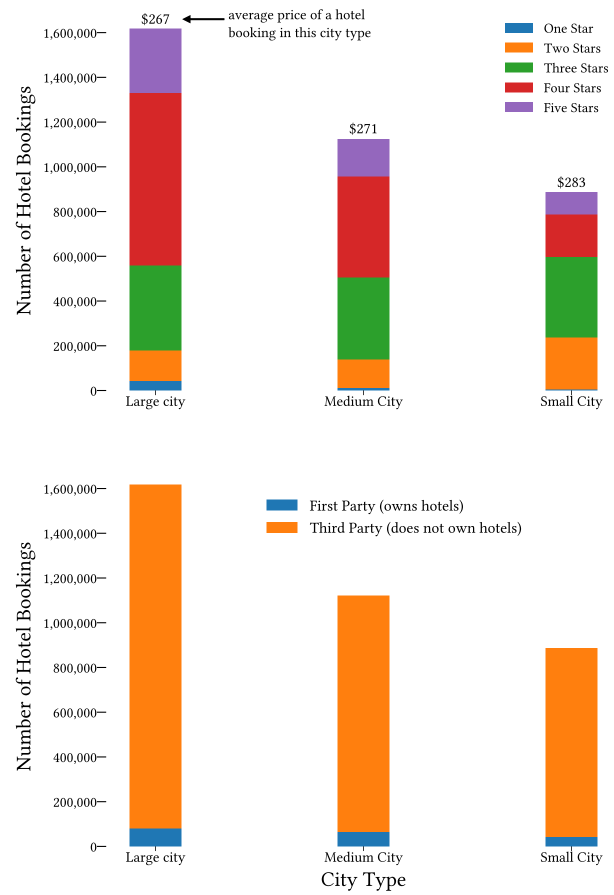
Seller	F_A	F_B	M_A	M_B	S_A	S_B	Y_A	Y_B
First Party 1	2.26	0.93	2.89	1.76	3.87	0.82	1.36	1.96
First Party 2	1.15	1.21	3.23	1.27	1.78	1.20	0.67	1.22
First Party 3	1.80	0.93	3.47	1.88	3.73	0.93	3.40	2.41
First Party 4	1.64	0.63	1.20	0.99	3.15	1.26	0.94	1.57
First Party 5	1.30	0.59	2.13	1.00	1.49	0.50	1.13	1.16
First Party 6	2.75	0.85	3.44	1.98	4.90	1.19	1.61	2.65
First Party 7	8.78	1.17	8.76	3.46	7.39	2.57	4.63	6.89
First Party 8	1.19	0.73	3.20	1.23	1.58	0.78	6.52	2.96
First Party 9	2.32	0.79	2.91	1.71	4.40	1.10	1.41	2.26
Third Party 1	5.75	1.46	5.37	3.76	6.15	2.56	2.01	4.97
Third Party 2	2.63	0.80	3.63	2.15	4.30	1.16	3.61	3.06
Third Party 3	2.55	0.82	3.52	1.79	3.86	0.96	2.10	2.57
Third Party 4	3.00	0.83	3.96	2.22	4.27	1.64	2.63	3.14
Third Party 5	0.16	0.42	0.25	0.28	0.04	0.19	0.23	0.25
Third Party 6	0.76	0.34	0.58	1.80	1.15	0.21	0.02	0.53
Third Party 7	0.33	0.42	0.72	0.39	0.19	0.26	0.58	0.43
Third Party 8	5.95	2.17	6.04	2.32	7.13	1.71	0.72	4.70
Third Party 9	2.30	0.79	2.87	1.67	4.26	1.08	1.36	2.14
Third Party 10	3.82	1.20	4.72	2.23	5.01	1.23	1.94	3.66
Third Party 11	0.66	0.29	0.60	1.41	0.96	0.20	0.02	0.44
Third Party 12	4.35	1.09	4.20	2.43	4.93	1.38	3.21	3.17
Third Party 13	5.07	1.45	5.25	3.72	6.19	2.46	2.98	4.81
Third Party 14	3.68	1.15	4.37	2.75	5.01	1.78	2.49	4.05
Third Party 15	2.84	0.90	3.46	2.02	4.96	1.23	1.66	2.74
Third Party 16	2.62	0.98	3.47	1.83	4.02	1.03	2.19	2.90
Third Party 17	1.65	0.61	1.21	0.95	3.14	1.30	0.94	1.57
Third Party 18	3.44	1.03	4.22	2.41	5.18	1.35	1.94	3.28
Third Party 19	1.50	0.69	1.54	0.72	3.74	0.78	1.53	1.16
Third Party 20	2.37	0.77	2.92	1.73	4.38	1.06	1.39	2.32
Third Party 21	1.45	0.69	2.21	1.26	2.08	0.66	0.78	1.46
Third Party 22	5.08	1.45	5.09	3.66	6.17	2.44	2.58	4.78
Third Party 23	1.83	0.68	2.96	1.46	2.96	0.82	1.88	1.77
Third Party 24	3.35	1.04	4.16	2.37	4.77	1.33	2.73	3.60
Third Party 25	2.64	1.12	3.81	2.12	4.34	1.25	2.26	3.13

(b) *Expected Dollar Loss* $\mathcal{M}_H$: Mean loss in dollars for an expanded subset of sellers and profiles with reference profile M_B. The range of impacts are notably smaller than with $\mathcal{M}_F$. First Party 1230 gives relatively consistent prices while Third Party 1015 ranges from -$ 0.30 to $ 2.86.

Seller	F_A	F_B	M_A	S_A	S_B	Y_A	Y_B	C
First Party 1002	0.29	0.07	0.31	0.26	0.11	0.26	0.52	0.31
First Party 1069	0.24	0.15	0.18	0.24	0.13	0.20	0.14	0.27
First Party 1092	0.04	0.03	0.03	0.07	0.04	0.13	0.35	0.12
First Party 1229	0.11	0.14	0.20	0.10	0.16	0.11	0.11	0.24
First Party 1230	0.05	0.05	0.04	0.05	0.11	0.08	0.06	0.04
First Party 1255	0.02	0.07	0.11	0.21	0.48	0.11	0.77	0.38
First Party 1259	0.34	0.64	1.14	0.57	0.63	0.28	1.29	1.51
First Party 1278	0.16	1.01	0.18	0.38	0.33	0.16	0.57	0.81
First Party 1291	0.25	0.14	0.18	0.24	0.14	0.19	0.14	0.27
First Party 1298	0.25	0.14	0.18	0.24	0.14	0.20	0.14	0.27
First Party 1301	0.25	0.14	0.18	0.24	0.13	0.19	0.14	0.27
First Party 1310	0.24	0.14	0.18	0.24	0.13	0.19	0.14	0.27
First Party 1318	0.25	0.14	0.18	0.24	0.13	0.20	0.14	0.27
First Party 1338	0.24	0.14	0.18	0.24	0.14	0.19	0.14	0.27
First Party 1356	0.25	0.14	0.18	0.24	0.13	0.19	0.14	0.27
First Party 1361	0.25	0.14	0.18	0.24	0.14	0.19	0.14	0.27
First Party 1362	0.25	0.14	0.18	0.24	0.14	0.19	0.14	0.27
First Party 1364	0.25	0.15	0.18	0.24	0.14	0.20	0.14	0.27
First Party 1380	0.24	0.14	0.18	0.24	0.14	0.19	0.14	0.27
First Party 1391	0.24	0.14	0.18	0.24	0.13	0.19	0.14	0.27
First Party 1392	0.24	0.14	0.18	0.24	0.13	0.19	0.14	0.27
First Party 1396	0.24	0.14	0.18	0.24	0.14	0.19	0.14	0.27
First Party 1397	0.25	0.14	0.18	0.24	0.14	0.20	0.14	0.27
First Party 1408	0.24	0.14	0.18	0.24	0.14	0.19	0.14	0.27
Third Party 1001	0.01	0.23	0.11	-0.02	0.23	0.09	0.39	0.66
Third Party 1005	0.08	0.26	-0.13	-0.07	-0.27	0.03	0.34	0.21
Third Party 1007	0.25	0.18	0.30	0.18	0.20	0.24	0.19	0.40
Third Party 1009	0.29	0.02	0.36	0.31	0.26	0.29	0.12	0.57
Third Party 1010	0.36	0.23	0.49	0.07	0.13	0.33	0.08	1.22
Third Party 1011	0.02	-0.14	0.20	0.04	0.03	0.16	0.26	1.35
Third Party 1015	0.25	1.68	-0.30	-0.18	1.41	0.97	0.04	2.86
Third Party 1016	0.11	-0.05	-0.02	-0.05	-0.26	0.14	0.50	1.61
Third Party 1017	0.32	0.13	0.40	0.27	0.30	0.24	0.03	0.85
Third Party 1020	0.40	0.20	0.46	0.27	0.39	0.32	0.50	0.92



(a): Descriptive statistics for 1.1 million unique flight records broken down by route type. The top plot breaks down each bar by the number of stops, while the bottom plot breaks down each bar by seller type (first or a third party). We observe a skew in the number of records towards the large-large hub routes, and the price shown on top of the bar shows that these are also more competitive, with prices increasing as airport size decreases.



(b): Descriptive statistics for 3.6 million unique hotel records broken down by city type. The top plot breaks down each bar by the number of stars, while the bottom plot breaks down each bar by seller type (first or a third party). We observe a skew in the number of records towards large cities, though it is not as dramatic as the flight skew. Large city hotel bookings are also more competitive, with prices increasing as city size decreases.

Figure 9: Descriptive statistics for flight market data (a) and hotel market data (b). In the top and bottom plots, the y-axis shows the number of records.