

**APPROXIMATING PURCHASE PROPENSITIES AND RESERVATION PRICES
FROM BROAD CONSUMER TRACKING***BY BENJAMIN REED SHILLER¹*Brandeis University, Waltham, U.S.A.*

A consumer's web-browsing history, now readily available, may be much more useful than demographics for both targeting advertisements and personalizing prices. Using a method that combines economic modeling and machine learning methods, I find a striking difference. Personalizing prices based on web-browsing histories increases profits by 12.99%. Using demographics alone to personalize prices raises profits by only 0.25%, suggesting the percent profit gain from personalized pricing has increased 50-fold. I then investigate whether regulations intended to prevent price gouging increase aggregate consumer surplus. Two feasible regulations considered offer at best modest improvements.

1. INTRODUCTION

Consumers are now tracked on the web by thousands of tracking and telecom companies.² These tracking data may be sold for the purpose of targeting advertisements and personalizing prices. In this article, I investigate whether web-browsing data improve predictions of purchase propensities, and whether they increase profits from personalized pricing compared to the scenario where only demographics are available for such targeting.

Historically, advertisers and sellers have had a limited ability to target advertisements and prices due to the nature of advertising transmission and the high costs of acquiring data beyond basic demographics. For example, television advertisers would place ads on shows that catered primarily to particular demographic groups, but individual viewers of a given show could not be separately targeted and shown different ads. Similarly, firms might offer discounts off the regular price to seniors, women, students, or other easily verifiable demographic groups. In recent years, however, robust markets for large data sets on individual behavior, popularly referred to as "big data," have developed. These data can be used to form a hedonic estimate of individuals' purchase propensities and reservation prices. Thus, on the Internet, ads and prices can be targeted to individuals instead of broad demographic groups.

Concerns related to data collection and targeting, such as privacy concerns or concerns about the equity of personalized pricing, are influencing policy. For example, the European Union recently introduced substantial data use regulations, the General Data Protection Regulation (GDPR), which requires Web sites to inform consumers of the specific uses of their data (e.g., personalizing prices) and to obtain their consent. In the United States, President Obama used his address to the Federal Trade Commission (FTC) in January 2015 to express his intent to introduce new legislation, a *Consumer Privacy Bill of Rights*, which would institute a

*Manuscript received October 2017; revised July 2019.

¹ I am very thankful for the comments and advice of Kathryn Graddy, Shane Greenstein, Phil Leslie, Imke Reimers, Katja Seim, and Joel Waldfogel, and for exceptionally useful feedback from the editor and two anonymous referees. Finally, I would like to thank Steven Knowles for imparting some of his IT-specific knowledge. Please address correspondence to: Benjamin Shiller, Department of Economics, Brandeis University, 415 South St., Waltham, MA 02453, U.S.A. Phone: 781-736-5205. E-mail: shiller@brandeis.edu.

² The Economist (2014) notes that over 1,300 firms are tracking consumers at the 100 most popular Web sites. Telecom companies may also sell browsing data: <https://tinyurl.com/mfup9ph>.

framework for promoting transparent use of data that is limited by consumers' consent.³ Excessive privacy invasions and personalized pricing were identified as two major motivating concerns in a White House Report that same month (Office of the President, 2015). Bergemann et al. (2015) validates these concerns by proving that there are a wide range of potential impacts of improved price discrimination on consumer welfare, ranging from eliminating consumer surplus entirely to passing the entire surplus gained from price discrimination to consumers. Intuitively, firms may extract a greater share of surplus when setting personalized prices close to reservation prices. However, consumer surplus can rise if previously unserved customers are offered lower prices when prices are personalized.⁴ This article empirically investigates the impacts of price personalization.

The closest article on behavioral targeting using web-browsing data is Goldfarb and Tucker (2011), which investigates the impact of a law limiting web tracking on stated purchase intentions among banner advertising viewers. They find a 65% reduction in advertising effectiveness.⁵ In this article, using a revealed preference approach, I also find evidence suggesting web-browsing histories are very effective for identifying avid consumers.

The prior literature on personalized pricing, beginning with Rossi et al. (1995, 1996) and Chintagunta et al. (2005), among many others, typically considers personalized pricing based on past purchase history of the same product.^{6,7} The basic idea is intuitive: If a consumer buys a product frequently, or was previously willing to pay a high price, he or she likely has a high reservation price for the product and can be charged higher prices in future interactions. One can thus use a small set of variables that are intuitively excellent predictors of willingness to pay to set prices.⁸ The models described in their papers were designed for contexts like pharmacies and grocery stores, where their methods are widely used today to generate personalized coupons, a less conspicuous form of personalized pricing. A concurrent paper, Dube and Misra (2017), also incorporates machine learning techniques and validates counterfactual predictions of the profit gained from personalized pricing via a field experiment. Their paper is complementary, although it has a somewhat different focus: It investigates the impact of price discrimination in business-to-business transactions, uses a much smaller set of explanatory variables (excluding online browsing histories), and uses an estimation algorithm too burdensome to be applied to data sets with very large numbers of explanatory variables.⁹ Finally, other recent papers have examined the impact of telematics-based monitoring on price discrimination in insurance markets (Reimers and Shiller, 2019) and digital disintermediation on negotiated prices (Peukert and Reimers, 2018).

This article presents a method for estimating purchase propensities and optimal individual-level prices from big observational data sets. The method is then used, in one context, to estimate individual purchase propensities and aggregate profits from personalized pricing if nearly 5,000 web-browsing variables are available. Of course, these analyses cannot, by themselves, prove that price discrimination has recently become more effective. Therefore, I compare these results

³ <https://tinyurl.com/y8w9nr7h>. Subsequent proposed bills have not yet been enacted, as of August 2019.

⁴ Robinson (1934) notes that price discrimination can increase aggregate surplus only if sales rise. If sales remain the same (or fall), then aggregate surplus falls because some goods are (inevitably) inefficiently allocated to consumers that have a lower willingness to pay than some unserved consumers offered higher prices.

⁵ Others have noted that measuring online advertising effectiveness is inherently difficult (Lewis and Rao, 2015; Johnson et al., 2016).

⁶ Hannak et al. (2014) and Mikians et al. (2012) found some evidence of personalized pricing online but did not study its effectiveness. Rossi et al. (1995, 1996) and Waldfogel (2015) do investigate personalized prices but do not use detailed web-browsing histories as explanatory variables.

⁷ Personalized marketing, including pricing, is often referred to in the marketing literature as "customer addressability."

⁸ Acquisti and Varian (2005) and Villas-Boas (2004) suggest that strategic consumers may alter their buying behavior to mask their level of interest in the product in order to receive lower prices in later periods. If they do so, it may limit the explanatory power of past purchases.

⁹ Dube and Misra (2017) show that accounting for the uncertainty in individualized parameter estimates (using hierarchical Bayesian methods) slightly increases the effectiveness of personalized pricing, due to Jensen's inequality.

to the analogous outcomes when only demographics—which have long been available—can be used for such targeting.

I employ a two-step estimation procedure. First, I use a least absolute shrinkage and selection operator (LASSO) penalized logit model to estimate individuals' purchase probabilities at observed, nonpersonalized prices. Next, I estimate individual-specific utility parameters in a multinomial logit framework to match the predicted purchase propensities from the first step. Note this method estimates heterogeneous preferences using variation across individuals, and parameters are identified even if only a single market (time period/location) is observed. Due to limitations of available data (lack of price variation, specific tier choice unobserved), not all parameters are identified without additional moment conditions, so I introduce a supply-side optimal pricing condition and an aggregate tier choice moment condition to address this issue.

The estimated individual-specific model parameters govern the relationship between prices and expected individual-level demand (e.g., the probability a particular consumer chooses each of Netflix's products). The slope of expected demand reflects the level of statistical uncertainty when predicting a consumer's choice. Given estimated individual-level demand, I calculate individual purchase propensities at observed prices, simulate counterfactual prices and profits under personalized pricing, and evaluate possible regulatory interventions.

Netflix provides an auspicious context for study. First, analyzing these questions requires individual-level data on both web-browsing histories and *all* purchases of a particular item, information that rarely appears together in data sets available to academic researchers.¹⁰ However, Netflix subscription status can easily be imputed from web-browsing histories, implying that all needed data reside in available online browsing data sets. Second, because purchases occur online, Netflix could implement personalized pricing based on web data. Third, because interactions with consumers take place online, Netflix could preempt consumer resentment by framing personalized pricing as automatically applied customized coupons—a strategy currently employed by some large online retailers—because coupons are already widely used and tolerated by consumers (Venkatesan and Farris, 2012).

Estimates from the first step are used to investigate the effectiveness of browsing data for targeting advertisements. I find that using web browsing histories to predict which consumers subscribe yields a higher true positive rate, and lower false positive rate, in a sample of 5,000 households held out from the estimation sample. Web-browsing data are thus more effective than demographics for finding consumers that are reasonably likely to make a purchase.

Simulations further reveal that incorporating web-browsing behaviors substantially raises the amount by which personalizing markups raises profits. Profits are only 0.25% higher if using demographics alone to personalize markups, but 12.99% higher if using all data. These findings suggest web-browsing data make personalized pricing substantially more appealing to firms.

In principle, price discrimination can raise both firm profits and consumer welfare (Bergermann et al., 2015). However, aggregate consumer surplus is found to be lower when markups are personalized, suggesting regulations may be appealing. I simulate outcomes under two different feasible types of regulation to investigate whether price regulations might yield a scenario where both firms and consumers are better off than they are under uniform pricing. In both cases, I find that firms charge somewhat higher prices to less avid consumers to satisfy regulations, preventing regulations from meaningfully raising consumer welfare.

The remainder of the article is organized as follows. Section 2 describes the context and industry background. Section 3 describes the data. Section 4 presents the model, and Section 5 provides estimation details. The main results of the article are then presented in Section 6. A brief conclusion then discusses concerns over perceived unfairness of personalized prices and explains how firms have circumvented these concerns and begun personalizing prices.

¹⁰ These data are, however, easily accessible by the firm.

2. BACKGROUND

Netflix, a DVD rentals-by-mail provider, was very popular in the year studied, 2006. Over the course of the year, 11.57 million U.S. households subscribed at some point (Netflix, 2006). This implies that about 16.7% of Internet-connected households consumed Netflix during 2006.¹¹

Netflix was differentiated from competing alternatives in at least three important ways. First, consumers' main alternative was to travel to a brick-and-mortar store (incurring a travel cost), instead of having movies arrive at their residences.¹² Second, by operating a few large warehouses outside of cities on cheaper land, Netflix was able to stock a much larger variety of videos than local brick-and-mortar rental outlets. Third, Netflix developed a well-regarded recommendation algorithm, which reduced consumers' search costs on the platform. A recent study (Gomez-Uribe and Hunt, 2016) suggests that 80% of viewing choices are attributable to the recommendation algorithm, whereas only 20% of viewing is attributable to consumer search. Given these advantages, Netflix is expected to have had some pricing power, at least during the period studied.

Netflix's subscription plans can be broken into two categories. Unlimited plans allow consumers to receive an unlimited number of DVDs by mail each month, but restrict the number of DVDs in a consumer's possession at one time. Limited plans set both a maximum number of DVDs the consumer can possess at one time and the maximum number sent in one month.

In 2006, there were seven plans to choose from. Three plans were limited. Consumers could receive one DVD per month for \$3.99 monthly; two DVDs per month, one at a time, for \$5.99; or four per month, two at a time, for \$11.99. The unlimited plan rates, for one to four DVDs at a time, were priced at \$9.99, \$14.99, \$17.99, and \$23.99, respectively.¹³ None of the plans allowed video streaming, since Netflix did not launch that service until 2007 (Netflix, 2006).

Key statistics for later analyses are the marginal costs of each plan. The marginal costs for the unlimited plans were estimated using industry statistics and expert guidance. They are assumed to equal \$6.28 for the one DVD at-a-time plan, \$9.43 for the two DVDs at-a-time plan, and \$11.32 for the three DVDs at-a-time plan.¹⁴

3. DATA

The data for this study were obtained from ComScore, through the Wharton Research Data Services (WRDS) interface. The data contain, for a large panel of computer users, demographic variables and the following variables for each Web site visit: the top-level domain name, time visit initiated and duration of visit, number of pages viewed on that Web site, the referring

¹¹ The total number of U.S. households in 2006, according to Census.gov, was 114.384 million (<http://www.census.gov/hhes/families/data/households.html>). About 60.6% were Internet-connected, according to linear interpolation from the respective numbers of connected homes in 2003 and 2007, according to the CPS Computer and Internet Use supplements. $11.57 / (0.606 \times 114.384) \times 100 \approx 16.7$.

¹² Blockbuster did offer a DVD-by-mail service starting in 2004 (<https://tinyurl.com/yc9fmpwx>). However, the service was marginalized because it competed with the core business that had invested in local brick-and-mortar outlets. The program only grew after incorporating in-store exchanges starting in November 2006. Subscriptions increased quickly, reaching two million in total by January 2007 (Netflix, 2006).

¹³ A very small number of buyers were observed paying \$16.99 per month for the three DVDs at-a-time unlimited plans. These observations were interspersed over time, suggesting it was not due to a change in the posted price.

¹⁴ A former Netflix employee recalled that the marginal costs of each plan were roughly proportional to the plan prices (i.e., the marginal cost for plan j approximately equaled $x \times P_j$, where x is a constant). I further assume that the marginal cost of a plan is unchanging and thus equal to the average variable cost. With this assumption, one can find x by dividing total variable costs by revenues. According to Netflix's financial statement, the costs of subscription and fulfillment, a rough approximation to total variable costs, were 62.9% of revenues, implying $x = 0.629$. Subscription and fulfillment include costs of postage, packaging, cost of content (DVDs), receiving and inspecting returned DVDs, and customer service. See Netflix (2006) for further details. The implied markup over variable cost is $\frac{1}{0.629} = 0.59$. The markups implied by financial reports from neighboring years were similar but slightly lower (0.53 in 2007, 0.46 in 2005).

Web site, and details on any transactions.¹⁵ For further details on this data set, refer to previous research using this data set (Moe and Fader, 2004; Montgomery et al., 2004; Huang et al., 2009).

Netflix subscription status can be imputed in these data. For a small sample of computer users observed purchasing Netflix on the tracked computer during 2006, subscription status is known. For the rest, I assume that a computer user is a subscriber if and only if he or she *averaged* more than two subpage views within Netflix's Web site per visit. The reasoning behind this rule is that subscribers have reason to visit more subpages within Netflix.com to search for movies, visit their queues, rate movies, and so forth. Nonsubscribers cannot access as many pages because they cannot sign in. According to this rule, 15.75% of households in the sample subscribe. This figure is within a single percentage point of the estimated share of U.S. Internet-connected households subscribing, found in Section 2. This small difference may be attributed to approximation errors in this latter estimate and ComScore's sampling methods.

For a small subset (a few hundred) of consumers, their Netflix transactions are recorded in the ComScore data, and hence their choice of Netflix tier (one, two, or three DVDs at a time) is observed. I use their tier choices to infer the aggregate share choosing each Netflix tier, assuming this group is representative of the larger population of Netflix subscribers.¹⁶ Specifically, the aggregate share choosing each tier is inferred by multiplying the fraction of consumers choosing any of Netflix's offerings (in the entire sample) by the fraction of consumers selecting each tier among the sample of consumers whose specific tier choice is recorded (in the small sample of consumers, who all subscribed).

I derived several web behavior variables from the raw data. These include the percent of a computer user's visits to all Web sites that occur at each time of day and on each day of the week. Time of day is broken into five categories: early morning (midnight to 6 a.m.), mid morning (6 a.m. to 9 a.m.), late morning (9 a.m. to noon), afternoon (noon to 5 p.m.), and evening (5 p.m. to midnight).

I then cleaned the data by removing Web sites associated with malware, third-party cookies, video rental chains, and pornography, leaving 4,600 popular Web sites to calculate additional variables.^{17,18} The total number of visits to all Web sites and to each single Web site were computed for each computer user.

The cross-sectional data set resulting from the above steps contains Netflix subscription status and a large number of variables for each of the 61,312 computer users.¹⁹ These variables can be classified into three types: standard demographics, basic web behavior, and detailed web behavior. Variables classified as standard demographics were race/ethnicity, children (Y/N), household income ranges, oldest household member's age range, household size ranges, census region (North, South, East, West), and area and population density of their ZIP code tabulation area from the 2010 Census. Variables classified as basic web behavior include the percent of online browsing by time of day and by day of week, and a broadband indicator. Summary

¹⁵ In our correspondence, ComScore representatives stated that demographics were captured for individual household members as they complete "a detailed opt-in process to participate," for which they were incentivized.

¹⁶ In unreported tests, I confirmed that this sample had a similar predicted propensity to subscribe to Netflix (according to Equation 10 using the full data set), as subscribers whose exact tier choice was unobserved, which might suggest the sample is akin to a random sample. However, it is not possible to test explicitly whether this sample of consumers is random for the purposes of the model, that is, it is not possible to verify whether this sample had the same propensity to consume higher quality tiers as the population at large. If the shares selecting each tier are atypical in this sample, it will bias estimates of the marginal intrinsic utilities for higher quality tiers, in the model described in Section 4.

¹⁷ yoyo.org provides a user-supplied list of some Web sites of dubious nature. Merging this list with the ComScore data reveals that such Web sites tend to have very high (≥ 0.9) or very low (≤ 0.1) rates of visits that were referred visits from another Web site, relative to sites not on the list, and rarely appear on Quantcast's top 10,000 Web site rankings. Web Sites were removed from the data accordingly, dropping sites with low or high rates referred to or not appearing in Quantcast's top 10,000. Manual inspection revealed that these rules were very effective in screening out dubious Web sites. In addition, video rental Web sites were dropped.

¹⁸ Pornography might contain valuable information, but might also require listing Web site names in publication that might offend some readers.

¹⁹ ComScore's data set was a rolling panel. Computers not observed for the full year were dropped. A couple hundred computer users with missing demographic information were also dropped.

TABLE 1
SUMMARY STATISTICS FOR HOUSEHOLD DEMOGRAPHICS AND BASIC BROWSING HABITS

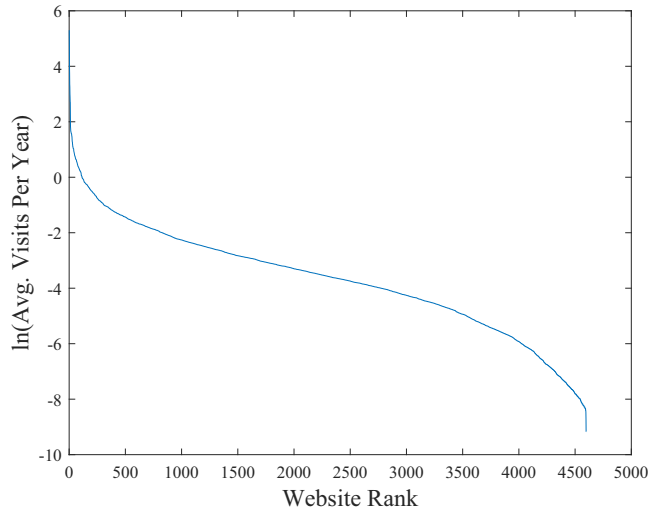
Household Demographics		Household Demographics			
I(Children)		Number of Residents			
Yes	59.5%	One Resident	13.4%		
		Two	32.9%		
Age (Eldest)		Three	21.2%		
Group 1 (18–20)	0.3%	Four	18.2%		
Group 2 (21–24)	2.0%	Five	9.8%		
Group 3 (25–29)	4.1%	Six or More	4.6%		
Group 4 (30–34)	14.9%				
Group 5 (35–39)	9.1%	Zipcode	Mean	St dev	
Group 6 (40–44)	19.7%	Area (Square Miles)	70.4	162.3	
Group 7 (45–49)	12.1%	Population Density	3,161.4	7,995.5	
Group 8 (50–54)	12.1%				
Group 9 (55–59)	8.7%				
Group 10 (60–64)	6.8%				
Group 11 (65+)	10.4%				
		Browsing Habits			
		I(Broadband)			
Race		Yes	78.5%		
Caucasian	94.5%				
Black	4.0 %				
Asian	1.1%	Timing of Internet Use	Mean	StDev	
Other	0.4%	Early Morning (Midnight to 6 a.m.)	8.7%	14.4	
		Mid Morning (6 a.m. to 9 a.m.)	11.7%	13.5	
I(Hispanic)		Late Morning (9 a.m. to Noon)	21.1%	13.9	
Yes	20.3%	Afternoon (Noon to 5 p.m.)	34.9%	16.1	
		Evening (6 p.m. to Midnight)	23.6%	26.5	
Income (in Thousands)					
Group 1 (<15)	9.8%	Monday	15%	4.0	
Group 2 (15–24.9)	6.5%	Tuesday	15.6%	4.8	
Group 3 (25–34.9)	10.9%	Wednesday	15.6%	4.6	
Group 4 (35–49.9)	20.3%	Thursday	15.2%	4.7	
Group 5 (50–74.9)	26.0%	Friday	14.6%	4.4	
Group 6 (75–99.9)	12.2%	Saturday	11.9%	7.0	
Group 7 (>100)	14.3%	Sunday	12.1%	7.2	

NOTES: The table reports summary statistics for the demographic and basic browsing behavior variables. The percent of households in the sample in the respective grouping is reported for binary grouping variables. For other variables that may vary over time for a given household, the mean and standard deviation are reported.

statistics for the demographic and basic browsing variables are shown in Table 1. Note the substantial variation in demographics that could be used to personalize prices.

Variables classified as detailed web behavior indicate number of visits to a particular Web site, with one variable for each of the 4,600 Web sites, as well as the household’s total Web site visits and transactions on the web during 2006. The yearly frequency of visits to individual Web sites, averaged across consumers, are shown in Figure 1. The most popular Web site in 2006, msn.com, was visited about 200 times a year by the average person, slightly more than once every other day. However, visits decline quickly with rank. The 50th ranked Web site (usps.com) was visited by the average customer about 2.2 times per year, partly because patrons visit it infrequently, and partly because 61 % of households in the sample never visited the site in 2006. Subsequent sections analyze whether such variations in behaviors across consumers, presumably indicative of preferences and habits, are useful for segmenting consumers and personalizing prices.

Prior to estimation, each explanatory variable is normalized to have a mean of zero and a standard deviation of one. This is common when using machine learning techniques that employ regularization.



NOTES: The figure shows the log frequency of visits to each Web site over the year, averaged across consumers. The Web sites are ordered according to this measure on the horizontal axis.

FIGURE 1

FREQUENCY OF VISITS ACROSS WEB SITES [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

The data were randomly split into two samples of individuals: a training sample of 56,312 individuals and a holdout sample of the remaining 5,000 individuals. The first, an estimation sample, is used for estimating model parameters. The model's fit is then tested using the sample of 5,000 consumers held out from estimation.

4. MODEL AND IDENTIFICATION

This section describes the strategy for estimating individual-level *expected* demand for Netflix's vertically differentiated products. The model presented is designed for the context studied. In other contexts, with richer data, one may be able to estimate a more flexible model.

In the chosen context (Netflix in 2006), prices did not vary across consumers or time. This alleviates endogeneity concerns, because prices cannot be correlated with changes in unobserved product attributes if prices do not change at all. Hence, the typical concern that endogenous prices will bias the price sensitivity parameter does not apply in this context. However, without price variation, price sensitivities are not identified from the demand-side model alone. Furthermore, without data on each individual's choice of quality tier, individual-specific marginal utilities for higher tiers are not identified.

In order to address these issues, I introduce a supply-side model and aggregate tier choice moment conditions (and model restrictions) that are used to augment the demand-side model. The augmented estimation model assumes that the researcher has *ex ante* knowledge of marginal costs. The supply-side model yields an optimal-pricing first-order condition used to estimate consumers' mean price sensitivity (assuming marginal costs are known). The augmented model also assumes that consumers agree on the marginal intrinsic utilities provided by higher-quality tiers. Aggregate tier choice shares are used to estimate the mean marginal intrinsic utilities for higher-quality tiers when individual-level tier choice is not consistently observed. Details are below.

4.1. Demand. The aim of the model is to estimate the relationship between prices and expected demand (i.e., probability of purchase) at the individual-consumer level. As in McManus (2008), a heterogeneous agent logit framework is used to model demand for vertically

differentiated products.²⁰ Specifically, consumer i 's conditional indirect utility for product j equals

$$(1) \quad u_{ij} = \alpha P_j + v_i + \delta_j + \sigma_\epsilon \epsilon_{ij} = \bar{u}_{ij} + \sigma_\epsilon \epsilon_{ij},$$

where P_j denotes product j 's price, and α and v_i denote price sensitivity and individual i 's intrinsic valuation for the product line, respectively.²¹ δ_j represents the marginal intrinsic valuation for higher tiers (δ_1 , corresponding to the lowest quality tier, is assumed to be equal zero, because it is not separately identified from the mean of v_i). ϵ_{ij} is an i.i.d. error term following the type 1 extreme value distribution.²² Finally, σ_ϵ denotes the standard deviation of the error term. The mean utility of the outside good is normalized to zero: $u_{i0} = 0 + \sigma_\epsilon \epsilon_{i0}$.

In the model, the parameters v_i , δ_j , and α represent estimable components of consumer preferences. Together, they determine consumer i 's *expected* reservation price, based on information available to the firm. The error term (ϵ_{ij}) reflects remaining uncertainty. As the error's scale (σ_ϵ) declines, prediction of which product provides highest utility (product choice) improves (i.e., product choice probabilities move closer to the extremes [0,1]). Analogously, as the scale of all remaining parameters (v_i , δ_j , α) increases, holding error scale (σ_ϵ) fixed, predictions of consumers' choices improve.²³ In practice, the latter (scale of remaining parameters [v_i , δ_j , α]) determines predictive ability because the error scale (σ_ϵ) is typically normalized (to one) for the model to be identified. Following convention, I subsequently assume $\sigma_\epsilon = 1$. Hence, if predicted-choice probabilities are near the extremes, zero and one, then the model will reflect these strong predictions by estimating a large scale for observable components (v_i , α , δ_j). This would imply that the components of an individual consumer's demand inferred from their browsing history are large compared to the unobserved components of their demand, implying that precise targeting is feasible.

The probability consumer i chooses product j is

$$(2) \quad s_{ij}(v_i, \alpha, \delta, P) = \frac{\exp(\alpha P_j + v_i + \delta_j)}{1 + \sum_{k \in J} \exp(\alpha P_k + v_i + \delta_k)},$$

where J denotes the set of inside products, and δ and P denote the product-tier vectors $[0, \delta_2, \delta_3]$ and $[P_1, P_2, P_3]$, respectively.

Correspondingly, the probability that consumer i chooses one of the inside products (Netflix's products), as opposed to the outside good (no Netflix product), is

$$(3) \quad s_{ij \neq 0}(v_i, \alpha, \delta, P) = 1 - s_{i0}(v_i, \alpha, \delta, P) = 1 - \frac{1}{1 + \sum_{k \in J} \exp(\alpha P_k + v_i + \delta_k)}.$$

²⁰ Similar models are often referred to as "random coefficient" models. However, coefficients are not random in this context; the model explicitly estimates separate coefficients for each observed individual.

²¹ In other contexts, a model that interacts the consumer preference parameter v_i with tier quality δ_j may be more appropriate. However, in this context, I found consumers with large subscription probabilities were not substantially more likely to choose higher quality tiers (in contrast to a priori expectations under a multiplicative model), in a small sample of consumers whose specific tier choice was observed.

²² I assume ϵ follows the type 1 extreme value distribution with location parameter equal to the negative of Euler's constant and scale parameter equal to one.

²³ Suppose that on estimable parameters (excluding error ϵ), product j is preferred to product k (i.e., $\alpha P_j + v_i + \delta_j > \alpha P_k + v_i + \delta_k$). Product j is chosen if $\alpha P_j + v_i + \delta_j + \epsilon_{ij} > \alpha P_k + v_i + \delta_k + \epsilon_{ik}$, which can be arranged to yield $\epsilon_{ik} < \alpha P_j + \delta_j - \alpha P_k - \delta_k + \epsilon_{ij}$. Thus, the probability that product j is preferred to product k is $F(\alpha P_j + \delta_j - \alpha P_k - \delta_k + \epsilon_{ij})$, where $F()$ denotes the cumulative distribution function (CDF) of error term ϵ . Increasing the scale of estimable parameters by multiplying them by the same constant (> 1) increases the positive difference between $(\alpha P_j + \delta_j)$ and $(\alpha P_k + \delta_k)$, which increases $F(\alpha P_j + \delta_j - \alpha P_k - \delta_k + \epsilon_{ij})$ (i.e., probability product j is preferred to product k). Hence, increasing the scale of estimable parameters v_i , α , and δ_j raises the probability of choosing the product providing higher utility based on estimable measures toward one, and reduces the probability that the other product(s) are chosen toward zero, implying better predictions.

The demand model is used to construct two sets of moment conditions: (1) (machine learning) estimated probabilities each individual i subscribes less the corresponding multinomial logit model's predictions ($\hat{s}_{ij \neq 0}(X_i) - s_{ij \neq 0}(v_i, \alpha, \delta, P)$) and (2) aggregate product tier shares less model predictions ($\hat{s}_j - s_j, \forall j$, where $s_j = \int s_{ij}(v_i, \alpha, \delta, P) f(v_i) dv_i$).

4.2. Supply. The supply-side model closely follows supply-side models in typical random coefficient models, with one important difference. Typically, price sensitivity is identified from the demand side, and the supply-side optimal-pricing first-order condition is used to recover estimates of marginal costs. However, because prices do not vary in the studied context, mean price sensitivity is not identified from the demand side. Instead, ex ante information on marginal cost is assumed (see Section 2), and the supply side is used to estimate mean price sensitivity.

The equation for profit can be simplified based on conversations with the former vice president of marketing at Netflix. He stated that Netflix's prices followed a simple pricing rule: Each tier had roughly the same percent markup over marginal cost. Specifically, $P_j(\theta) = (1 + \theta)c_j$, where c_j is the marginal cost of tier j and θ is a markup parameter. Information in Netflix's annual report implies $\theta = 0.59$, and the marginal costs for the three tiers are thus \$6.28, \$9.43, and \$11.32.²⁴ Under this pricing strategy, the expression for profit is

$$(4) \quad \pi = \sum_{j \in J} (P_j(\theta) - c_j) M s_j - \Gamma = \sum_{j \in J} \theta c_j M s_j - \Gamma,$$

where M is the mass of consumers, Γ is the fixed cost, and s_j is the aggregate share selecting tier j .

The first-order condition yields

$$(5) \quad \frac{d\pi}{d\theta} = \sum_{j \in J} c_j \left(s_j + \theta \frac{ds_j}{d\theta} \right) = 0,$$

where $\frac{ds_j}{d\theta}$ equals

$$(6) \quad \frac{ds_j}{d\theta} = \int \frac{ds_{ij}(v_i, \alpha, \delta, P(\theta))}{d\theta} f(v_i) dv_i,$$

where $P(\theta)$ denotes the price vector as a function of markup ($P(\theta) = (1 + \theta) \times [c_1, c_2, c_3]$), $f(v_i)$ denotes the density of the individual-specific parameters, and $\frac{ds_{ij}(v_i, \alpha, \delta, P(\theta))}{d\theta}$ follows the logit-model framework.²⁵ Specifically,

$$(7) \quad \frac{ds_{ij}(v_i, \alpha, \delta, P(\theta))}{d\theta} = \alpha s_{ij}(v_i, \alpha, \delta, P(\theta)) \left(c_j - \sum_{k \in J} c_k s_{ik}(v_i, \alpha, \delta, P(\theta)) \right).$$

The third moment condition used in estimation is the derivative of profits with respect to markup ($\frac{d\pi}{d\theta}$), from Equation (5). It should equal zero, assuming that the firm is choosing a markup to maximize profits.

²⁴ See Section 2 for details.

²⁵ $s_{ij}(v_i, \alpha, \delta, P(\theta)) = \frac{\exp(\alpha(1+\theta)c_j + v_i + \delta_j)}{1 + \sum_{k \in J} \exp(\alpha(1+\theta)c_k + v_i + \delta_k)}$. Hence, $\frac{ds_{ij}(v_i, \alpha, \delta, P(\theta))}{d\theta} = \frac{\alpha c_j \exp(\alpha(1+\theta)c_j + v_i + \delta_j)}{1 + \sum_{k \in J} \exp(\alpha(1+\theta)c_k + v_i + \delta_k)} - \frac{\exp(\alpha(1+\theta)c_j + v_i + \delta_j)}{(1 + \sum_{k \in J} \exp(\alpha(1+\theta)c_k + v_i + \delta_k))^2} (\sum_{k \in J} \alpha c_k \exp(\alpha(1+\theta)c_k + v_i + \delta_k))$. $\rightarrow \frac{ds_{ij}(v_i, \alpha, \delta, P(\theta))}{d\theta} = \alpha c_j s_{ij}(v_i, \alpha, \delta, P(\theta)) - s_{ij}(v_i, \alpha, \delta, P(\theta)) (\sum_{k \in J} \alpha c_k s_{ik}(v_i, \alpha, \delta, P(\theta)))$. $\rightarrow \frac{ds_{ij}(v_i, \alpha, \delta, P(\theta))}{d\theta} = \alpha s_{ij}(v_i, \alpha, \delta, P(\theta)) (c_j - \sum_{k \in J} c_k s_{ik}(v_i, \alpha, \delta, P(\theta)))$.

4.3. *Moment Conditions.* To recap, the aim is to estimate preference parameters (v_i, α, δ_j) . There are three sets of moment conditions used in estimation in the typical case where the data are less than ideal.

The objective function used to estimate these parameters is:

$$(8) \quad \min_{\alpha, \delta_1, \dots, \delta_J, v_1, \dots, v_N} \left(\begin{aligned} & \left(\sum_{i=1}^N (\hat{s}_{ij \neq 0}(X_i) - s_{ij \neq 0}(v_i, \alpha, \delta, P))^2 \right. \\ & \left. + \sum_{j \in J} (\hat{s}_j - \int s_{ij}(v_i, \alpha, \delta, P(\theta)) f(v_i) dv_i)^2 \right) \\ & \left. + \left(\sum_{j \in J} c_j \left(s_j + \theta \frac{ds_j}{d\theta} \right) \right)^2 \right). \end{aligned} \right).$$

The first set of moment conditions is the difference between estimates of individuals' probabilities of subscribing to any tier of Netflix's products less the corresponding multinomial logit-model predictions. The second is the difference between the aggregate share selecting tier j (\hat{s}_j) and the corresponding model prediction (s_j), $\forall j$, where $s_j = \int s_{ij}(v_i, \alpha, \delta, P(\theta)) f(v_i) dv_i$. The last moment condition is the derivative of profits with respect to markup. Estimated demand parameters (v_i, α, δ) minimize the objective.²⁶

4.4. *Identification and Restrictions.* Note that the implied probability of subscribing ($s_{ij \neq 0}(v_i, \alpha, \delta, P)$) is monotonically increasing in v_i . Hence, conditional on the values of other parameters, v_i are identified by the first set of moment conditions, which equate the predicted probability each consumer subscribes (to any of the inside tiers) under the structural framework with the corresponding machine learning prediction from the first step.²⁷ In the model presented above, there is an analytic solution for v_i .²⁸ Note that identification of v_i does not require observations from multiple markets— v_i are identified by variation in subscription probabilities across individuals.

The incremental valuations for higher tiers (δ_2, δ_3) are identified by the second set of moment conditions, which match observed aggregate tier shares (\hat{s}_j) and model predictions ($s_j = \int s_{ij}(v_i, \alpha, \delta, P(\theta)) f(v_i) dv_i$). Note, a unique set of δ_j satisfy these moment conditions, because the predicted aggregate share choosing a given tier j is monotonically increasing in δ_j .

Next, note that information on tier choice probabilities under a single price schedule is not sufficient to identify all preference parameters (v_i, α, δ_j) . If one both (i) reduces the price sensitivity (α) by constant A , and (ii) adds a constant $A \times P_j$ to consumer i 's intrinsic utility for each inside product j ($v_i + \delta_j$), then the mean utility for each tier j (and hence choice probability for each tier) remains unchanged.²⁹

However, additional information on demand elasticities identifies α and $(v_i + \delta_j)$ separately. Note, if holding choice probabilities $s_{ij}(v_i, \alpha, \delta, P)$ fixed to satisfy the first two sets of moment conditions (i.e., holding fixed conditional indirect utility by adjusting $v_i + \delta_j$ to compensate for any changes in α), then demand elasticities depend on only the price sensitivity α . For example, in the logit model, an individual's own price elasticity equals $\left(\frac{ds_{ij}(v_i, \alpha, \delta, P)}{dP_j} \frac{P_j}{s_{ij}(v_i, \alpha, \delta, P)} \right) = \alpha P_j (1 - s_{ij}(v_i, \alpha, \delta, P))$, implying a monotonic relationship between α and demand elasticity. Thus, only one set of α and $(v_i + \delta_j)$ can rationalize both the choice probabilities and elasticities. Any added information on choice elasticities allows separate identification of α and $v_i + \delta_j$.

In our context prices do not vary. Hence, price sensitivity (α) and intrinsic values for the inside products ($v_i + \delta_j$) are not separately identified from the demand side. One can address this issue by using information from the supply side, assuming a constant markup. Assume that for any change in α the parameters $v_i + \delta_j$ are adjusted accordingly so that subscription probabilities (s_{ij}) are kept the same, and the first and second sets of moment condition remain

²⁶ The model is just identified. Hence, at the optimum, the objective function equals zero.

²⁷ See Lewbel (2019) and Matzkin (1992).

²⁸ When the individual-specific term v_i enters additively into the utility function, and the error is i.i.d. and follows the type 1 extreme value distribution, an analytic solution exists: $v_i = \ln \left(\frac{\hat{s}_{ij \neq 0}(X_i)}{(1 - \hat{s}_{i, j \neq 0}(X_i)) \sum_{k \in J} \exp(\alpha P_k + \delta_k)} \right)$.

²⁹ $(v_i + \delta_j + A \times P_j) + (\alpha - A)P_j = v_{ij} + \delta_j + \alpha P_j$.

satisfied. Then, larger negative values of α imply more elastic demand and thus lower optimal markup θ (over ex ante known marginal costs). Only one value of α implies observed markups are optimally set, satisfying the third moment condition ($d\pi/d\theta = 0$).³⁰

4.5. Estimation Routine. I search for parameter values minimizing the objective function stated above in Equation (8). In order to speed computation, I search over the representative price sensitivity (α) and marginal utilities for higher tiers (δ_j , for $j > 1$), nesting within a search for the individual-level parameters (v_i) equalizing the logit model predicted probabilities of selecting an inside product ($s_{ij \neq 0}(v_i, \alpha, \delta, P)$) and the corresponding estimates from the LASSO penalized logit model ($\hat{s}_{ij \neq 0}(X_i)$).³¹

5. ESTIMATION

The model is estimated in two steps. The first step—the main focus of this section—estimates purchase probabilities for individual consumers. In the second step, remaining model parameters are estimated as described in the preceding section.

A logistic regression with LASSO regularization is used to estimate the probability each consumer subscribes to any one of Netflix's services (one, two, or three DVDs at a time), as a function of individual-level observables (e.g., browsing data). The penalized likelihood function equals

$$(9) \quad L = \sum_i \ln(s_{ij \neq 0}(X_i) \times I(buy) + (1 - s_{ij \neq 0}(X_i)) \times (1 - I(buy))) - \lambda \sum_{k=1}^K |\beta_k|,$$

where $s_{ij \neq 0}(X_i)$ denotes the model-predicted probability of subscribing, $I(buy)$ indicates whether consumer i subscribes, λ is the LASSO penalty parameter, β_k denotes the value of parameter k , and

$$(10) \quad s_{ij \neq 0}(X_i) = \frac{\exp(\phi + X_i\beta)}{1 + \exp(\phi + X_i\beta)}.$$

The estimation routine searches for parameters (ϕ, β) to maximize the penalized likelihood for a given penalty parameter. The penalty parameter (λ) is estimated by maximizing the out-of-sample likelihood using twofold cross validation.

A set of 4,633 explanatory variables are considered. In all models, 18 demographic variables are considered, including (i) indicators for children, race, Hispanic ethnicity, census region (North, West, and South); (ii) ordinal group number (e.g., 18- to 20-year-olds are considered group number 1), the group number squared, and indicator for censored from above, for each of the following: age, income, and household-size groupings; and (iii) linear measures of the area and population density of the household's ZIP code tabulation area. Additionally, some models include the remaining 4,615 explanatory variables summarizing individuals' web browsing, including indicators for the browsing habits listed in Table 1, the intensity of web use (number site visits), and its square, and the number of visits at each of 4,600 Web sites. All explanatory variables are normalized to have mean 0 and standard deviation of 1, prior to estimation.

³⁰ Combining Equations (5), (6), and (7), and solving for α yields an analytic solution for the unique value of α satisfying the last (optimal pricing) moment condition: $\alpha = \frac{\sum_{j \in J} c_j s_j}{\theta \sum_{j \in J} (\frac{1}{N} \sum_{i=1}^N s_{ij} (c_j - \sum_{k \in J} c_k s_{ik}))}$.

³¹ An analytic solution for v_i exists: $\hat{s}_{ij \neq 0}(X_i) = 1 - \frac{1}{1 + \sum_{k \in J} \exp(\alpha P_k + v_i + \delta_k)} \rightarrow v_i = \ln\left(\frac{\hat{s}_{ij \neq 0}(X_i)}{(1 - \hat{s}_{ij \neq 0}(X_i)) \sum_{k \in J} \exp(\alpha P_k + \delta_k)}\right)$.

TABLE 2
ESTIMATION RESULTS: PREDICTION WHETHER INDIVIDUAL SUBSCRIBES

Order	Visits to:	Coef. Value	Δ Pr(Subscribe)
			(w. One Standard Deviation Increase in Visits)
1	gamefly.com	0.78	11.78
2	ameblo.jp	-0.48	-5.03
3	slysoft.com	0.36	4.80
4	audible.com	0.33	4.40
5	dvdfab.com	0.30	4.01
6	sutterhealth.org	0.30	3.96
7	4chan.org	0.29	3.84
8	jambase.com	0.29	3.83
9	imdb.com	0.28	3.72
10	houstonpress.com	0.27	3.49
11	kw.com	-0.26	-2.90
12	somethingawful.com	-0.25	-2.83
13	jacksonville.com	-0.24	-2.72
14	lacity.org	-0.24	-2.65
15	jalopyjournal.com	0.23	3.03
16	uhaul.com	0.23	2.99
17	smackjeeves.com	0.23	2.98
18	dailypress.com	-0.23	-2.58
19	sonlight-email.com	-0.22	-2.54
20	fairfaxcounty.gov	0.22	2.89
21	ganeshaspeaks.com	0.22	2.89
22	onstation.com	-0.22	-2.49
23	whig.com	0.21	2.74
24	techdirt.com	0.21	2.68
25	zylom.com	-0.21	-2.36
26	npr.org	0.20	2.64
27	baseballamerica.com	0.20	2.63
28	apunkachoice.com	0.20	2.63
29	elpais.com	0.20	2.61
30	amazon.com	0.20	2.56

NOTES: The table shows coefficient estimates from a logistic regression with LASSO regularization. The 30 variables with largest coefficients are reported. Prior to estimation, all variables are normalized, so the mean and standard deviation of visits across individuals are 0 and 1, respectively. Column (3) reports the coefficient estimates. Column (4) shows the corresponding percentage point increase in subscription probability arising from a one standard deviation increase in the respective variable, averaged across consumers.

A useful feature of LASSO models is that the procedure selects potentially meaningful explanatory variables, setting coefficients on other variables to zero.³² In the full model, a large number, 938 of the 4,633 considered explanatory variables, remain and have nonzero coefficients. Of these, only five are demographic variables, suggesting that the web-browsing data provide richer information on consumers’ tastes.

The top 30 variables in the full model, ordered by coefficient magnitude, are reported in Table 2. Because all variables are normalized prior to estimation, coefficient size provides a measure of the influence of the variable on estimated subscription probabilities. Note that the 30 most influential explanatory variables do not include any demographic variables, suggesting demographic variables provide little marginal information about subscription probabilities. Also note that the list does not include any basic browsing behaviors like timing or intensity of web use. Rather, a household’s tendency to visit particular Web sites seems to contain the most information about their affinity for Netflix.

The intuition linking Web site visits to an innate affinity for Netflix is obvious for some. GameFly (#1), Audible (#4), and Amazon (#30)—all positive predictors of Netflix subscription—

³² Testing revealed that regularization methods yielded a somewhat better out-of-sample fit, compared with stepwise models commonly used in the economics literature in the past.

indicate a preference for receiving products by mail, suggesting a higher valuation for Netflix, which at that time delivered DVDs by mail. 4chan.org (#7) is a site for anime enthusiasts interested in content that is typically available at Netflix—which operates large warehouses—but may not be available at the local brick-and-mortar outlets comprising Netflix’s competition in 2006. The top sites also include a site catering to consumers interested in technology startups (#26, techdirt) and a movie review Web site (#9, imdb.com). In order to address the concern that movie review Web sites might be complements for Netflix, raising concerns of reverse causality, the model is reestimated, excluding all Web sites categorized as related to “movies” or “TV” (including imdb.com, DVDfab.com, and slysoft.com), according to Alexa web rankings.³³ These results are reported in Subsection 6.4.

The impacts of the most influential variables are large. The last column of Table 2 shows the corresponding marginal change in subscription probability, on average across consumers, occurring when visits to the Web site increase by one standard deviation. For example, users who visit GameFly.com one standard deviation more often than average are 11.8 percentage points more likely to subscribe to Netflix at observed prices, implying a 75% ($11.8/16 \times 100$) higher probability of subscribing than the mean consumer, whose subscription probability is about 16%.

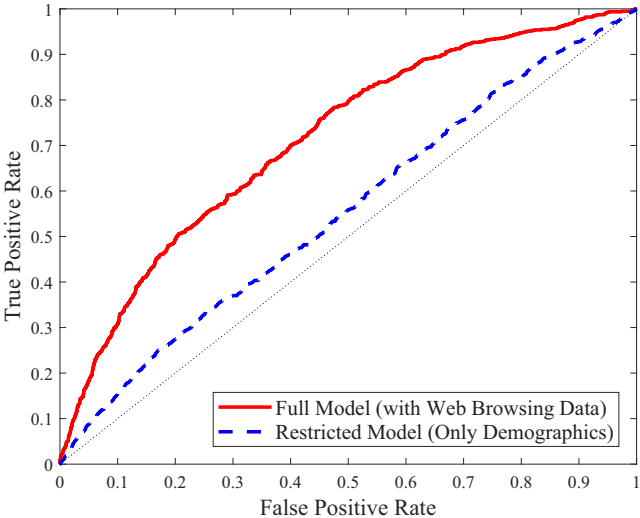
The logit demand model parameters (v_i, α, δ_j) are then estimated via generalized method of moments, as described in Section 4. Recall that estimated subscription probabilities $(s_{ij \neq 0}(X_i))$ from the first stage of the model, the LASSO penalized logit model described in this section, are an input into the second stage, which estimates the structural parameters (v_i, α, δ_j) . Thus, the structural model parameter estimates depend on the set of variables used to estimate subscription probabilities $s_{ij \neq 0}(X_i)$ in the first stage. I estimate model parameters (v_i, α, δ_j) twice, once using the estimated probability each consumer subscribes based on demographics alone, and a second time using the estimated probability each consumer subscribes when all variables (demographics and web-browsing data) are used. Profits and welfare under counterfactual pricing schemes can be simulated for each of these sets of parameter estimates, and are compared in Subsection 6.2.

6. RESULTS

6.1. Efficacy of Targeting. One can analyze how well various data types identify likely consumers using only the first stage of the model, which predicts the probability a consumer subscribes from a large set of variables without relying on the assumptions of the structural model. Several tests yield the same conclusion: Web-browsing data yield superior predictions of consumers’ choices.

I first employ a conventional method used to evaluate fit in the machine learning literature, the receiver-operating characteristic (ROC) plot (Zweig and Campbell, 1993). The ROC plot shows the true positive rate (share Netflix subscribers that are predicted to subscribe) against the false positive rate (share of nonsubscribers predicted to subscribe). The plot is created by varying an arbitrary threshold. For example, the true and false positive rates can be calculated using a 75% threshold, that is, assigning individuals as predicted subscribers if their subscription probability, according to the model, exceeds 75%. The ROC plot shows the true positive and false positive rates for an array of different thresholds values. A 45 degree line implies the model holds no predictive value. Curves farther above the 45 degree line imply better fit. The area under (the ROC) curve (AUC) yields an intuitively ordinal measure of fit that can be used to compare various models. An AUC value of 0.5 implies the model has no predictive value. AUC values closer to the maximum possible value, one, imply better fit.

³³ In some contexts, a subset of sites’ visits could correlate with purchase because they are complements, not because they are a priori useful for segmenting consumers. When designing a pricing experiment, practitioners should use caution, and if possible, utilize browsing data from a period prior to the consumers’ purchase decisions.



NOTES: The figure shows the receiver-operating characteristic (ROC) plot for the model predicting individual subscription probabilities based on demographic variables, and the model based on all variables. The ROC plot is constructed by varying the threshold the predicted probability of subscribing must surpass in order for an individual to be classified as a *predicted* subscriber, and for each threshold, calculating the true positive rates (share Netflix subscribers that are predicted to subscribe) and false positive rates (share of nonsubscribers predicted to subscribe).

FIGURE 2

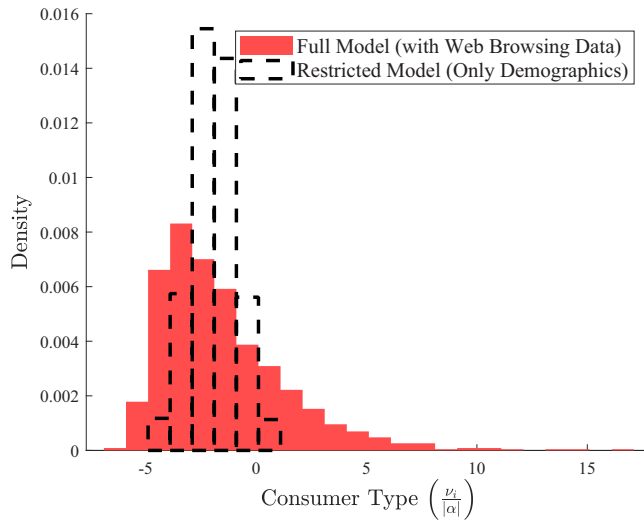
MODEL FIT—ROC PLOT [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

The ROC plot generated using the sample held out from estimation is shown in Figure 2. Visual inspection confirms that the curve for the model predicting subscription using all variables, including browsing histories, lies much farther above the 45 degree line, compared with the curve for the model using only demographic variables, suggesting the model based on all variables yields a much better fit. The AUC of the full model, 0.71, is substantially larger than the AUC of the demographics-based model, 0.55, confirming that the model with web-browsing variables yields much better predictions of which consumers subscribe.

Figure 3 illustrates the extent of heterogeneity that can be captured by the different variable sets. The consumer type, defined as $\frac{v_i}{|\alpha|}$, is shown on the x -axis. It represents the expected intrinsic utility for the lowest-quality tier, normalized by the price sensitivity. Note, the consumer type omits the unobserved error ϵ_{i1} ; A consumer's actual valuation for the lowest quality tier equals $\frac{v_i + \epsilon_{i1}}{\alpha}$. Also note that the consumer type ($\frac{v_i}{|\alpha|}$) can be negative, just as intrinsic utilities in classic random coefficient models are negative when purchase probabilities are somewhat small—only those consumers that have a large enough unobserved utility error (ϵ) choose to purchase the respective product. Note that the range of consumer type values is much wider when all information, including web-browsing histories, is used to predict which consumers are most likely to purchase.

6.2. Personalized Pricing Counterfactuals. This section simulates counterfactual environments in which Netflix implements personalized pricing, proper second-degree price discrimination, or both. Specifically, optimal profits under each pricing scheme are simulated separately, first using demographics alone, and then using all variables to explain a consumer's willingness to pay.³⁴ They are then compared with simulated profits under the status quo environment, where Netflix employed constant percent-markup pricing.

³⁴ In principle, consumers' observed choices provide information on the ranges of ϵ_{ij} consistent with their choices. However, because the goal is to set prices to new consumers before their choices are observed, instead of conditioning price on past purchase history, this information is disregarded.



NOTES: This figure shows a histogram of consumer types $\frac{v_i}{|\alpha|}$ observable under the full model (which includes web-browsing histories) and the restricted model (which uses only demographics to predict purchase).

FIGURE 3

EXTENT OF HETEROGENEITY PREDICTABLE USING VARIOUS DATASETS [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

TABLE 3
SIMULATED CHANGES IN VARIOUS OUTCOMES RESULTING FROM PERSONALIZED MARKUPS

	Percent Change When Price Based on:	
	Demographics	All Variables
Total Profits	0.27% (0.01)	12.99% (0.23)
Subscribers	0.13% (0.00)	3.12% (0.03)
Sales (DVDs at-a-Time)	0.14% (0.00)	3.21% (0.05)
Aggregate Consumer Surplus	0.04% (0.00)	-0.50% (0.08)
Joint Surplus	0.05% (0.00)	0.91% (0.02)

NOTES: The table reports the changes in simulated outcomes when markups are personalized, compared to when they are not. Bootstrapped standard errors in parentheses.

Table 3 shows the percent increase in profits from personalizing markups (θ_i), where percent markup differs across individuals, but for a given individual is the same across products.³⁵ When all variables are used to personalize markups, profits increase 12.99%.³⁶ If personalized markups are based only on demographics, the increase in total profits is much less, 0.25%. The nascent availability of these data thus increases the likelihood that firms will implement price

³⁵ Percentages instead of absolute profits were reported because simulated variable profits in the status quo case depend on the demand estimates, which can vary slightly depending on which set of variables were used in estimation. In practice, the two status quo profit estimates were quite close, within about half of a percent of each other.

³⁶ In this calculation, variable costs are defined as the “cost of revenues” reported in Netflix’s 2006 Annual Report (Netflix, 2006). The “operating expenses” in the 2006 financial report are assumed to be fixed costs. These definitions imply the variable costs were about \$627 million, and the fixed costs were about \$305 million. Revenues in 2006 were about \$997 million, implying variable profits were about \$370 million, and total profits were about \$65 million. Multiplying the percent change in variable profits by 370/65 yields the percent change in total profits.

TABLE 4
PERCENT INCREASE IN PROFITS, RELATIVE TO NONPERSONALIZED LINEAR PRICING

	Nonpersonalized	Personalized
Linear Pricing	N/A	12.99% (0.23)
Second-Degree PD	22.48% (0.17)	36.86% (0.15)

NOTES: Linear pricing denotes a consistent percent markup (θ) over cost across all three tiers. Second-degree PD (price discrimination) denotes that the percent markup (θ_j) may differ across the tiers. *Personalized* denotes consumers are charged different prices based on their browsing histories and demographics. Bootstrapped standard errors in parentheses.

discrimination, because adding web-browsing data substantially increases the amount by which personalizing markups raises profits.

Personalizing markups reduces aggregate consumer surplus by 0.50%.³⁷ However, most consumers receive lower prices when prices are personalized, and hence the majority are slightly better off.

Next, as a comparison, I consider second-degree price discrimination, where all consumers face the same prices, but there is a separate markup for each product tier (θ_j). Suppose that when a consumer arrives, preferences are private (i.e., the firm has no information on individual-specific preferences). Under second-degree price discrimination, the markups for each tier are designed so that consumers self-select the tier with an incremental price about equal to their willingness to pay for the incremental quality.³⁸ Despite the fact that all consumers face the same price schedule, consumers end up paying different prices per unit of quality and the firm’s expected profits earned from a randomly arriving consumer increase.

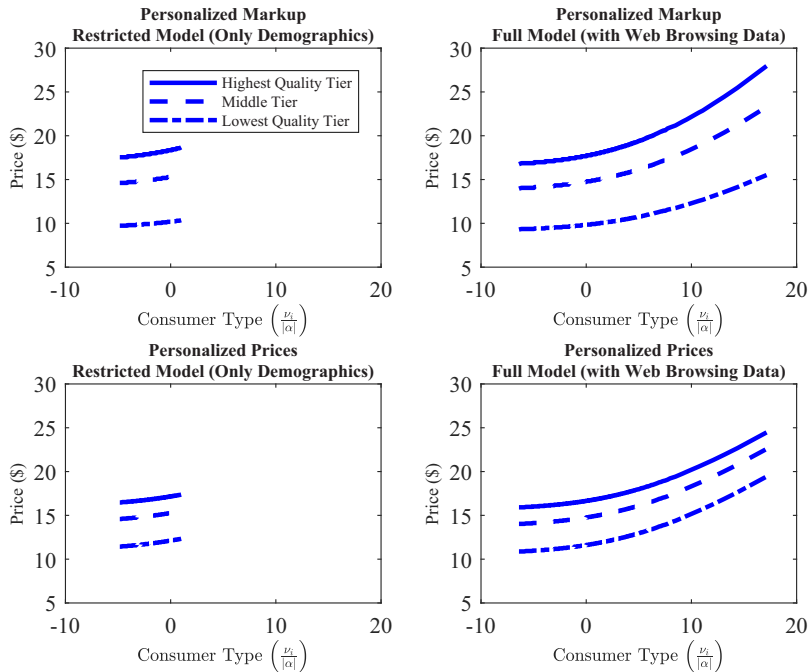
Personalization and second-degree price discrimination are not mutually exclusive, and combining the strategies by setting a separate markup (θ_{ij}) for every combination of consumer and product tier may increase profits further. When a consumer arrives, the firm may have some information on his or her preferences (e.g., based on their web-browsing history). However, a consumer’s exact preferences remain private information. Based on available information and the remaining uncertainty, the firm can form an expected distribution of parameters governing a consumer’s preferences. From the firm’s perspective, an arriving consumer represents a group of heterogeneous consumers with private preferences, from which the arriving consumer’s actual realized preferences are drawn. Just as classic second-degree price discrimination can increase expected profits from an anonymous consumer, combining personalization with the mechanics of second-degree price discrimination can raise profits when consumers retain some private information about their preferences.

The results are shown in Table 4. Changing from the status quo case (constant percent markup over cost, θ) to second-degree price discrimination (separate markup θ_j for each tier) increases profits by 22.48%. Switching instead to personalized markups (θ_i) raises profits by 12.99%. Combining the strategies by setting a separate markup (θ_{ij}) for each combination of consumer i and tier j raises profits by 36.8%.

Second-degree price discrimination raises profits by raising the price of the lowest quality tier ($\approx 18\%$) and reducing the price of the highest tier ($\approx 6\%$). Increasing the price of the lowest tier reduces profits from that tier, but reduces the price differences between higher tiers and

³⁷ Percent change in consumer surplus equals $\frac{CS_{personalized} - CS_{status\ quo}}{CS_{status\ quo}} \times 100$, where under the logit modeling assumptions, consumer surplus for a given pricing scheme equals $\frac{\sum_i \ln(1 + \sum_{k \in J} \exp(\alpha P_k + v_i + \delta_j))}{\alpha}$.

³⁸ In classic second-degree price discrimination theoretical models, if the firm chooses qualities for an arbitrarily large set of vertically differentiated goods, each consumer pays exactly their full willingness to pay for marginal quality. With a small discrete set of goods, some consumers may pay less than their full valuation for marginal quality.



NOTES: In each panel, personalized prices for each tier are plotted against consumer type. The top left panel shows the range of prices (for each tier) across consumers when the percent markup over cost is personalized using demographic information. The top right panel shows prices across consumers when the percent markup is personalized using the full set of variables. The bottom left panel shows the range of prices across consumers when the firm personalizes the prices of each tier (personalized second-degree price discrimination) using demographics, and the bottom right shows the corresponding range of prices when prices of each tier are personalized using the full set of available variables.

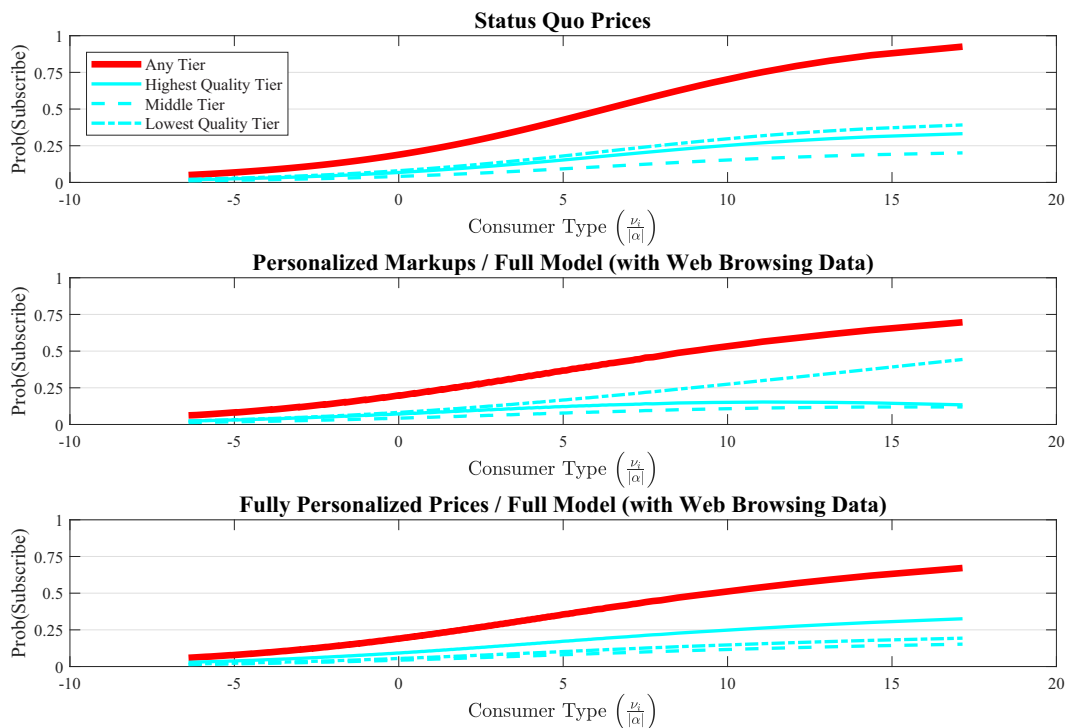
FIGURE 4

RANGE OF PERSONALIZED PRICES ACROSS CONSUMERS, FOR ALL THREE TIERS [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

the lowest tier and thus loosens the incentive compatibility constraints.³⁹ More consumers are willing to pay the resulting smaller incremental price to acquire a higher quality tier, shifting some consumers to higher tiers that have larger absolute markups. Overall, profits increase. Similar patterns occur when personalizing separate markups for each tier (personalized second-degree price discrimination).

Using the full set of variables to personalize markups substantially increases the range of prices charged to different individuals for the same product. Figure 4 shows price against consumer type, separately by tier, in four scenarios. The first scenario, appearing in the top left panel, shows that if the percent markup is personalized using demographic variables, the highest offered prices are about 6% higher than the lowest offered prices, for each respective tier. If markups are personalized using the full set of variables (top right panel), the highest prices are 67% higher than the lowest prices offered, for each tier, and 55% higher than the prices offered when markups are not personalized. The 99th percentile consumer is offered a price about 16% higher than the nonpersonalized price, and the 95th percentile consumer is offered a price that is 5% higher. The 84th percentile consumer faces essentially the same prices whether or not markups are personalized. This is not that surprising, because under status quo prices, only the 16% of consumers with highest valuations actually subscribed to Netflix. All consumers

³⁹ The derivative of variable profit from tier j with respect to its price equals: $\int_i \frac{\partial \pi_{ij}}{\partial P_j} df_i = \int_i s_j + \alpha s_{ij}(1 - s_{ij})(P_j - c_j) df_i$. It is negative for the first tier. The derivative of the profit for tier j , for $j > 1$, with respect to the price of the first tier is $\int_i \frac{\partial \pi_{jk}}{\partial P_j} df_i = - \int_i \alpha s_{ij} s_{ik} (P_k - c_k) df_i$, which is weakly positive because $\alpha < 0$.



NOTES: All three panels show subscription probabilities against consumer type, when the full set of variables (demographics and web-browsing histories) are used to estimate the model. The top panel shows the tier choice probabilities for different consumers under status quo prices. The middle panel shows tier choice probabilities under personalized percent markups. The bottom panel shows tier choice probabilities if personalizing prices for each of the three tiers separately.

FIGURE 5

COUNTERFACTUAL TIER CHOICE PROBABILITIES [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

with lower predicted valuations are offered lower prices when markups are personalized, some receiving discounts as large as 7%.

The bottom panels of Figure 4 show the ranges of prices offered to different consumers when prices are fully personalized, that is, when the firm personalizes prices separately for each tier using demographics (left panel) and using all variable (right panel). Note that compared to the case where the firm only personalizes markups (top right panel), when the firm personalizes the prices of all tiers separately (bottom right panel) the price of the lowest quality tier is noticeably higher for consumers with high predicted valuations, to encourage captive consumers to select a higher-quality tier yielding a larger absolute markup. This is confirmed in Figure 5. Raising prices for the lowest-quality tier encourages high-value consumers to switch to the higher quality tiers, which have larger absolute markups. Even though raising the price of the lowest-quality tier results in those consumers who are (wrongly) predicted to have high valuations leaving Netflix, the gain from encouraging others to switch to more profitable tiers more than offsets these losses.

6.3. *Regulations.* In principle, price regulations could yield personalized prices that raise both profits and aggregate consumer surplus, compared with the status quo case, a uniform markup charged to all consumers. This can be demonstrated with a simple example. Suppose personalized markups are offered to consumers whose optimal personalized markup is below the flat markup observed, and all other consumers continue to receive the same markup as before. No consumers are offered higher prices. Thus, no one is worse off. However, consumers offered

a lower price are better off than before, so aggregate consumer surplus rises. Personalizing prices to a subset of consumers, as opposed to no consumers, increases the firm's profit as well.

In practice, however, mutually beneficial regulations may be elusive. In a world with price discrimination, regulators may not be apprised of the uniform pricing markup, or consumer demand in general. Hence, the simplest regulation—explicitly setting a price/markup ceiling—may be challenging to implement well.

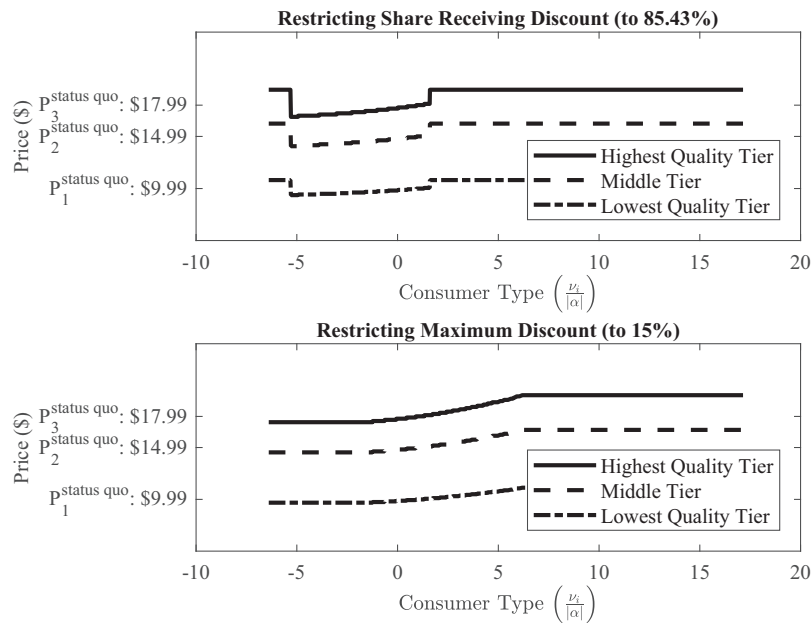
Instead, two more feasible price regulations are considered. The first type of feasible regulation will be referred to as a discount penetration regulation, as it limits the percentage of consumers receiving a discount off a list price chosen by the firm. If the discount penetration ceiling is set to 100% (no regulation), then the firm sets a high list price and offers each consumer a personalized discount. This is equivalent to fully personalized pricing. At the other extreme, a discount penetration ceiling of 0% prohibits the firm from offering a discount to any consumer and is therefore equivalent to banning price discrimination. If regulators set the discount penetration rate somewhere in between, at, say, 85%, then the firm could offer as many as 85% of consumers a discount off the list prices.

The impacts of price regulations capping the discount penetration at $n\%$ are simulated. The simulation consists of an inner step and an outer loop. The outer loop performs a grid search over possible markup ceilings (list prices) to find the markup ceiling that maximizes profits. The inner step selects which consumers to offer discounts off the list prices in order to maximize profits. First, one calculates the expected profits for each consumer if that consumer was offered the list prices. Consumers are then ordered according to the difference between the expected profits they generate at the list prices and at the optimal personalized markup. In order to satisfy the price regulation while minimizing the impact on profits, the $(100 - n)$ percent of consumers with lowest difference are offered the list prices. The remaining $n\%$ receive personalized discounts. Summing expected profits across consumers yields the total expected profits for a given guess at the markup ceiling (list prices) in the outer loop. The markup ceiling maximizing profits (in the outer loop) is chosen.

This entire simulation process is repeated for a range of discount penetration regulation limits, ranging from $n = 0\%$ (same markups for all consumers) to $n = 100\%$ (no regulation). The impact of these regulations on the firm is intuitive. Profits are monotonically declining in the strength of price regulations (inverse of discount penetration rate).

The simulations reveal that prohibiting price discrimination entirely maximizes consumer surplus. However, consumer surplus can be higher absent regulation than with limited regulation. Reducing the percent of consumers allowed to receive personalized discounts from 100% (no regulation) to 97% reduces consumer surplus by 0.13%. For discount penetration cap regulations between 1% and 96%, consumers in aggregate are better off than they are under full price discrimination, but worse off than they are under uniform pricing. Thus, although it is theoretically possible for less stringent regulations to induce personalized prices that increase consumer surplus, such an outcome is not realized. Consumer surplus is maximized when price discrimination is banned entirely.

In order to understand why consumers are unable to benefit from the less stringent regulations, consider the impact at the individual consumer level. Consider the case where the discount penetration cap is set 85.43% (i.e., the percent of consumers whose profit-maximizing personalized markup is below the uniform markup observed empirically). The profit-maximizing markup ceiling (determining list prices) is 72%, which is higher than the observed uniform markup of 59%. This higher markup aims to extract surplus from avid customers. However, to adhere to the regulations' quota, the firm must offer this markup to 14.53% of consumers. The list price markup (72%) thus must be applied to some consumers who would be offered lower personalized prices absent regulations. Not surprisingly, the optimal personalized markups for some of these consumers are close to the list markup. However, the firm also charges list prices to some consumers who are unlikely to buy at any price, consumers with much lower optimal-personalized markups, because these consumers generate little profit for the firm regardless. The firm is better off using these low value consumers to satisfy the regulations' list



NOTES: In each panel, personalized prices for each tier are plotted against consumer type. The top panel shows prices under a regulation limiting the fraction of consumers offered a discount off some list (highest) price. The bottom panel shows prices under a regulation that limits the depth of the discount off the list price.

FIGURE 6

PRICES OFFERED ACROSS CONSUMERS UNDER PRICING REGULATIONS

price quota, allowing the firm to personalize discounts to other (mid value) consumers. See the top panel of Figure 6. Notice that limited regulations result in some consumers paying higher prices than they would absent regulations, and some paying higher prices than they would if price discrimination were banned entirely. Thus, some consumers are made worse off by this limited regulation, regardless of whether the alternative is full regulation (banning price discrimination), or no regulation.

Next, I investigate the impacts of a second feasible price regulation that limits the depth of the discount, that is, a maximum difference in prices across consumers, instead of limiting the fraction of consumers receiving a discount. In this scenario, the firm is free to set the maximum markup (price), and provide personalized discounts to any number of consumers. However, the depth of the discount cannot exceed some set threshold. In order to evaluate this type of regulation, I simulate outcomes under different discount depth limits between 5% and 50%, in increments of five.

Although outcomes are slightly different under a regulation that limits discount depth, the conclusions are essentially unchanged. Regulating markup personalization by setting a maximum discount depth weakly reduces firm profits. When the discount percentage is capped at 10%, aggregate consumers surplus is slightly higher, 0.09% higher. However, for other limits considered, aggregate consumer surplus is lower, and in most instances lower than it is without regulation. Again, the reason is that consumers predicted to have low willingness to pay actually face higher prices under the regulation, and are made worse off. Referring to the bottom panel of Figure 6, one can see personalized prices when the discount depth is capped at 15%. Compared to the first feasible price regulation (limiting the fraction of consumers receiving a discount), the least avid consumers (consumer types ($\frac{v_i}{\alpha}$) below -6) receive lower price offers. However, the lowest price offered to anyone when discount depth is capped at 15% is 3.3% larger than the lowest price offered under the first regulation, and 4% larger than lowest price offered absent regulation.

TABLE 5
ROBUSTNESS CHECKS

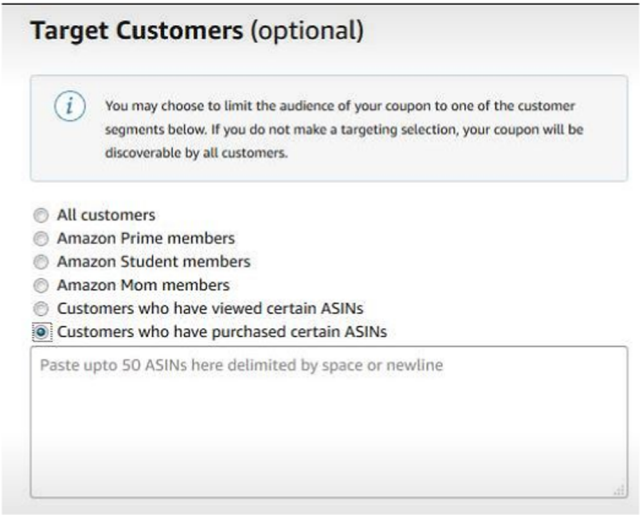
	Percent Increase in Profits When Price Personalization Based on:	
	Demographics	All Variables
Main Model	0.27% (0.01)	12.99% (0.23)
Robustness Checks		
Larger Assumed Markup	0.27% (0.01)	13.00% (0.29)
Excluding Movie and TV Sites	N/A	10.50% (0.44)
Excluding Extremes (1%)	0.07% (0.01)	6.92% (0.22)

NOTES: The first robustness check reestimates the model assuming that optimally set static prices are twice as high as observed prices, implying a much larger markup. The second robustness check reestimates the model excluding any site categorized as related to movies or TV, according to Alexa Web Rankings, from the set of Web sites used to predict subscription and affinity for Netflix. The last robustness check resimulates profits under both status quo and personalized markups, excluding the top and bottom half a percent of consumers with highest/lowest predicted affinity for Netflix. Bootstrapped standard errors in parentheses.

The firm's optimal behavior under these regulations is not necessarily surprising. Firms sacrifice profits from the least avid consumers to satisfy regulations, while preserving the ability to profit by raising prices to avid consumers. In fact, 42% of the gains from personalization are attributed to the 0.5% of consumers predicted to have the highest valuations for Netflix, 59% of the gains come from the 1% most avid, and 78% of the gains from the 5% most avid. Note that because only 16% of Internet-connected consumers subscribed, the top 0.5% of consumers account for over 3% of *subscribers*, and the top 5% of consumers account for nearly one third of actual subscribers. Still, these findings show that most of the profit gains from personalizing markups arise from increasing prices to the most avid consumers.

In summary, discriminatory pricing can in theory increase both producer surplus and aggregate consumer surplus. Regulators, therefore, might attempt to use limited regulations to reach such an outcome, instead of prohibiting personalized pricing entirely. However, simulations reveal that less stringent price regulations may either yield very modest gains to aggregate consumer surplus, or, in many cases, actually reduce consumer surplus. Moreover, these regulations have the unintended consequence of increasing prices for the least avid consumers, consumers that would otherwise be offered the largest discounts. Less stringent regulations are thus unappealing in some ways.

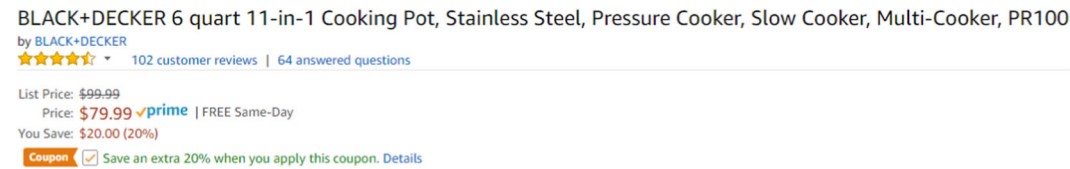
6.4. Robustness Checks. Table 5 shows that the increase in profits from personalized pricing is robust to some modeling assumptions. The first concern is that Netflix may have underpriced in the short run to grow the business, and hence the static optimal-pricing conditions may rely on a questionable assumption. However, I find that even if one assumes that the optimal prices were double the observed prices (implying a much higher markup), and reestimates the model, the increase in profits from price personalization is roughly the same, at least in percentage terms. The second concern is that movie review Web sites such as IMDB.com and RottenTomatoes.com might be complements for Netflix's products. If so, visiting them might not just indicate a higher intrinsic affinity for Netflix, but also indicate that a consumer already subscribes. Although this may be true, the impact of movie review Web sites on the main results is relatively small. Dropping all Web sites categorized by Alexa Web Rankings as related to movies or TV, along with Wikipedia, and rerunning the model lowers the percent gain from price personalization from 12.99% to 10.49%. Even with such sites dropped, personalized pricing based on browsing histories is much more effective than personalized pricing based



SOURCE: <https://tinyurl.com/y827fuzr>.

FIGURE 7

SELLER INTERFACE ALLOWING SELLERS TO PERSONALIZE COUPONS ON AMAZON [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]



NOTES: Screenshot taken on Amazon.com on September 21, 2018, while logged into a user’s account. In order to redeem the coupon, the consumer merely needs to click the box. At checkout, the price is listed at \$63.99, instead of the originally listed price of \$79.99.

FIGURE 8

EXAMPLE OF CLICKABLE COUPON ON AMAZON [COLOR FIGURE CAN BE VIEWED AT WILEYONLINELIBRARY.COM]

on only demographics. Finally, I repeat personalized pricing simulations excluding the most and least avid consumers, the top and bottom half percent, out of concern that functional form assumptions may cause imprecision in estimates at the extremes. Excluding these extreme consumers, the profit increase from personalized pricing is estimated to be 0.075% when only using demographics, and 6.9% when browsing data are used to personalize prices. Although the gain to personalized pricing falls when the extreme consumers are excluded, the main conclusion remains the same: Web-browsing data are much more effective than demographic data for personalizing prices.

7. DISCUSSION AND CONCLUSION

This article finds, in the context of Netflix, that web-browsing data substantially improve targeted advertising and increase profits from personalized pricing, relative to when only demographic data are available. Better ad targeting may benefit advertisers, firms, and even consumers, but the efficiency and equity effects of widespread personalized pricing are less well understood. Most textbooks espouse the efficiency of personalized pricing based on partial equilibrium analysis. However, I find that the benefits accrue to firms, not consumers in

aggregate, and feasible regulations allowing limited amounts of price discrimination do not appear to yield large benefits.⁴⁰

A related question is whether it is fair for consumers to pay different prices for the same product. Kahneman et al. (1986) finds personalized pricing was viewed as unfair by 91% of respondents. Yet, the prevalence of third-degree price discrimination (and coupons) suggests firms can profit by (effectively) offering different prices to different groups, and therefore consumers are willing to pay different prices than others for the exact same good.

Still, perceived fairness remains an important business consideration, and Amazon's multiple attempts at personalized pricing are an instructive example. In the year 2000, customers who discovered they were being offered different prices on Amazon reacted with fury.⁴¹ In response, Amazon refunded price differences and stopped personalizing prices. Over the last decade, a variety of firms began employing personalized pricing but were more careful than Amazon about framing. Personalized prices were not called their true name, but rather were labeled "customized coupons" or "personalized discounts." They are not coupons in the traditional sense; they are automatically applied when the consumer reaches the Web site and typically require no action on the customer's part (or sometimes just one click). They are merely called "coupons" or "discounts" to address concerns over perceived unfairness. In 2017, Amazon followed suit. Specifically, Amazon began allowing third-party sellers to offer coupons to consumers who had viewed or purchased certain products, designated by Amazon's internal product codes (ASINs).⁴² Figure 7 shows an example of the seller interface allowing targeted couponing. Using the coupon requires trivial effort from the targeted consumer, they merely click the coupon box, as shown in the example in Figure 8. Hence, Amazon allows third-party sellers to offer different prices to consumers based on their browsing histories at Amazon. It may not be widespread at Amazon (yet), but its existence and growing use comprises a conspicuous example of a major retailer embracing personalized pricing, but under a different name to address concerns over perceived unfairness.

REFERENCES

- ACQUISTI, A., AND H. R. VARIAN, "Conditioning Prices on Purchase History," *Marketing Science* 24(3) (2005), 367–81.
- BERGEMANN, D., B. BROOKS, AND S. MORRIS, "The Limits of Price Discrimination," *The American Economic Review* 105(3) (2015), 921–57.
- CHINTAGUNTA, P., J.-P. DUBE, AND K. GOH, "Beyond the Endogeneity Bias: The Effect of Unmeasured Brand Characteristics on Household-Level Brand Choice Models," *Management Science* 51(5) (2005), 832–49.
- CHOUDHARY, V., A. GHOSE, T. MUKHOPADHYAY, AND U. RAJAN, "Personalized Pricing and Quality Differentiation," *Management Science* 51(7) (2005), 1120–30.
- CORTS, K. S., "Third-Degree Price Discrimination in Oligopoly: All-Out Competition and Strategic Commitment," *The RAND Journal of Economics* 29(2) (1998), 306–23.
- DUBE, J.-P., AND S. MISRA, "Scalable Price Targeting," NBER Working Paper No. 23775, 2017.
- Economist, "Getting to Know You: Everything People Do Online Is Avidly Followed by Advertisers and Third-Party Trackers." *The Economist Magazine*, 2014.
- GOLDFARB, A., AND C. E. TUCKER, "Privacy Regulation and Online Advertising," *Management Science* 57(1) (2011), 57–71.
- GOMEZ-URIBE, C. A., AND N. HUNT, "The Netflix Recommender System: Algorithms, Business Value, and Innovation," *ACM Transactions on Management Information Systems (TMIS)* 6(4) (2016), 13:1–13:19.
- HANNAK, A., G. SOELLER, D. LAZER, A. MISLOVE, AND C. WILSON, "Measuring Price Discrimination and Steering on E-commerce Web Sites," In *Proceedings of the 2014 Conference on Internet Measurement Conference*, pp. 305–318. ACM (2014).

⁴⁰ When price discrimination is employed by multiple firms, the implications are more nuanced. In homogeneous-product oligopolistic markets (Spulber, 1979) and differentiated product markets (Thisse and Vives, 1988; Corts, 1998; Choudhary et al., 2005), personalized price discrimination does unilaterally raise profits, but employed jointly it may increase competition, reducing profits.

⁴¹ <https://tinyurl.com/y76k4mp7>

⁴² <https://www.skubana.com/amazon-coupon/>

- HUANG, P., N. H. LURIE, AND S. MITRA, "Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods," *Journal of Marketing* 73(2) (2009), 55–69.
- JOHNSON, G. A., R. A. LEWIS, AND D. H. REILEY, "When Less Is More: Data and Power in Advertising Experiments," *Marketing Science* 36(1) (2016), 43–53.
- KAHNEMAN, D., J. L. KNETSCH, AND R. THALER, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *The American Economic Review* 76(4) (1986), 728–41.
- LEWBEL, A., "The Identification Zoo—Meanings of Identification in Econometrics," *Journal of Economic Literature* 57(4) (2019), 835–903.
- LEWIS, R. A., AND J. M. RAO, "The Unfavorable Economics of Measuring the Returns to Advertising," *The Quarterly Journal of Economics* 130(4) (2015), 1941–73.
- MATZKIN, R. L., "Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models," *Econometrica: Journal of the Econometric Society* 60(2) (1992), 239–70.
- MCMANUS, B., "Nonlinear Pricing in an Oligopoly Market: The Case of Specialty Coffee," *The RAND Journal of Economics* 38(2) (2008), 512–32.
- MIKIAN, J., L. GYARMATI, V. ERRAMILLI, AND N. LAOUTARIS, "Detecting Price and Search Discrimination on the Internet," *Proceedings of the 11th ACM Workshop on Hot Topics in Networks* (2012), 79–84.
- MOE, W., AND P. S. FADER, "Dynamic Conversion Behavior at E-commerce Sites," *Management Science* 50(3) (2004), 326–35.
- MONTGOMERY, A. L., S. LI, K. SRINIVASAN, AND J. C. LIECHTY, "Modeling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science* 23(4) (2004), 579–95.
- Netflix, "Netflix Annual Report." <http://ir.netflix.com/annuals.cfm>, 2006.
- Office of the President, "Big Data and Differential Pricing," Washington, DC, 2015.
- PEUKERT, C., AND I. REIMERS, "Digital Disintermediation and Efficiency in the Market for Ideas," SSRN Working Paper 3110105, 2018.
- REIMERS, I., AND B. SHILLER, "The Impacts of Telematics on Competition and Consumer Behavior in Insurance," *The Journal of Law and Economics*, forthcoming (2019).
- ROBINSON, J., *The Economics of Imperfect Competition* (London: MacMillan and Co, Ltd, 1934).
- ROSSI, P. E., R. E. McCulloch, AND G. M. ALLENBY, "Hierarchical Modelling of Consumer Heterogeneity: An Application to Target Marketing," in C. Gatsonis, J. S. HODGES, R. E. Kass, and N. D. Singpurwalla, eds., *Case Studies in Bayesian Statistics*, Volume II (New York: Springer, 1995), 323–49.
- , "The Value of Purchase History Data in Target Marketing," *Marketing Science* 15(4) (1996), 321–40.
- SPULBER, D. F., "Non-Cooperative Equilibrium with Price Discriminating Firms," *Economics Letters* 4(3) (1979), 221–27.
- THISSE, J.-F., AND X. VIVES, "On the Strategic Choice of Spatial Price Policy," *The American Economic Review* 78(1) (1988), 122–37.
- VENKATESAN, R., AND P. W. FARRIS, "Measuring and Managing Returns from Retailer-Customized Coupon Campaigns," *Journal of marketing* 76(1) (2012), 76–94.
- VILLAS-BOAS, J. M., "Price Cycles in Markets with Customer Recognition," *RAND Journal of Economics* 35(3) (2004), 486–501.
- WALDFOGEL, J., "First Degree Price Discrimination Goes to School," *The Journal of Industrial Economics* 63(4) (2015), 569–97.
- ZWEIG, M. H., AND G. CAMPBELL, "Receiver-Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine," *Clinical Chemistry* 39(4) (1993), 561–77.