The Steering Incentives of Gatekeepers in the Telecommunications Industry*

Brian McManus   Aviv Nevo   Zachary Nolan
Jonathan W. Williams

May 9, 2023

Abstract

We study trade-offs faced by multiple-system operators (MSOs) when setting prices and quality for internet access and TV. As over-the-top video (OTT) improves, MSOs choose between accommodating OTT to share in the surplus it provides consumers, and steering consumers towards TV. To disentangle these incentives, we estimate a demand model using household panel data on subscription choices and internet usage. We find that if an MSO can price different types of internet usage, under many cost circumstances the MSO discounts OTT. Furthermore, the MSO's incentive to improve OTT quality is strengthened when it can price internet usage.

Keywords: Steering, Gatekeeper, Foreclosure, Bundling, Telecommunications Industry, Broadband Internet, Net Neutrality

JEL Codes: L11, L13, L96.

---

*This paper is a substantially revised version of a paper entitled "Steering Incentives of Platforms: Evidence from the Telecommunications Industry." We are grateful for the North American MSO that provided the data used in this paper. We thank John Asker, Greg Crawford, Shane Greenstein, Robin Lee, Jeff Prince, Tommaso Valletti and Ali Yurukoglu for comments on earlier versions of this paper, as well as participants in several seminars. McManus and Williams are grateful for support from NSF grant SES-1824293. Williams thanks Cable Television Laboratories, Inc., and its member organizations, for their support of this work. McManus: Department of Economics, University of North Carolina at Chapel Hill, mcmanusb@unc.edu; Nevo: Department of Economics and Wharton School, University of Pennsylvania, nevo@upenn.edu; Nolan: Department of Business Administration, University of Delaware, znolan@udel.edu; Williams: Department of Economics, University of North Carolina at Chapel Hill, jonwms@unc.edu. Nevo is currently on leave and serving as the Director of the Bureau of Economics at the FTC. The opinions expressed in this paper do not represent the views of the FTC or any individual Commissioner.

1 Introduction

Firms that sell internet access serve as gatekeepers to online content, including over-the-top video (OTT). Internet access providers typically also sell traditional TV service, to which OTT may be a direct substitute. This raises the concern that these multiple system operators (MSOs) may have an incentive to steer consumers toward their TV service by either increasing the price of internet service or degrading the access to and quality of OTT content. If MSOs were to take these actions, OTT providers and the consumers who use these increasingly popular services could be harmed. Concerns over harm to consumers and OTT providers are at the heart of the net neutrality debate, and they have been considered in public actions such as the FCC order issued after Charter's acquisition of Time Warner Cable.[1]

We study these issues by analyzing the steering incentives of MSOs. As OTT streaming services improve in quality, an MSO benefits through increased demand for its internet service. However, these benefits may be offset if consumers respond to OTT's improvement by substituting away from TV service.[2] The MSO therefore faces a trade-off between encouraging the growth of OTT so it may increase profit from an improved internet offering, and steering consumers away from OTT and towards its own video service.

To study this trade-off, we develop and estimate a model that captures the industry's salient features. An MSO offers a menu of standalone and bundled plans for internet and TV, and consumers subscribe to plans to access content. Given a subscription, consumers make internet and TV usage decisions to maximize utility from content exclusive to the internet and video content that can be accessed through a TV subscription or OTT. By modeling consumers' content usage choices, we can study consumers' responses to marginal prices for internet usage. These responses occur both at the intensive margin of usage decisions and at the extensive margin of plan choice. Our model, therefore, allows us to capture how usage prices can act as both steering instruments between internet and TV plans, and as metering tools to capture surplus from internet usage.

We estimate the model using household-level panel data from a North American MSO.

[1]For example, the order restricts Charter's ability to offer usage-based pricing out of concern that it will harm OTT. See https://docs.fcc.gov/public/attachments/FCC-16-59A1.pdf.

[2]Internet usage has grown steadily during recent years, largely driven by an increase in streaming video. About 60M U.S. households (46%) used a streaming video service in 2018, up from 44M in 2016 (comScore, 2018). Cisco, a major telecommunications and IT firm, estimates that 81% of North American internet usage was video during 2017, and this share will grow to 85% by 2022 (Cisco, 2018). The emergence and popularity of OTT services coincides with a trend in consumers dropping their TV service ("cord cutting") and instead consuming video through the internet. Between 2014 and 2017 in the U.S., the number of consumers who cut the cord grew from 3.1M to 14.1M (MarketWatch, 2018).

To capture the heterogeneous behavior that we observe across households in the data, we allow for heterogeneity in consumers' demand for overall internet usage, streaming video content, internet speed, and in their price sensitivity. We use a flexible fixed grid approach (Ackerberg, 2009; Fox et al., 2011, 2016; Malone et al., 2021a) to estimate the distribution of preference heterogeneity. An important feature of the data is the introduction of usage-based pricing (UBP) during the sample period. The UBP takes the form of a menu of multipart prices, each of which includes an access fee, an internet usage allowance, and an overage fee schedule. The policy's introduction helps us identify the parameters of our demand model.

The estimates allow us to quantify the MSO's incentives to steer consumers toward the bundle versus meter internet usage. We do this by solving for two usage prices—one for all internet usage and a second that applies only to OTT usage—as a function of the cost of providing internet and TV services. We find that the OTT-specific price can be positive or negative depending on costs. In scenarios that resemble current industry conditions in the US, with low TV margins and high internet margins, our estimates imply that the MSO prefers to set a lower price for OTT than other internet usage. Therefore, an evaluation of concerns about usage-related prices must separate the MSO's metering and steering incentives, or it may otherwise miss the strength and even the direction of the firm's steering incentive.

Our findings are relevant to the debate on MSOs' incentives under net neutrality. In particular, we find that the OTT-specific price is one example of a policy that can increase or decrease consumers' benefit from streaming video. We also study this issue directly by allowing an MSO to set the quality of OTT, which in our model is linked to the degree of substitution between TV and OTT. We show that the MSO may have an incentive to increase or decrease OTT quality, again depending on relative profit margins. Furthermore, we establish that, when the MSO uses UBP, there is a larger set of cost conditions under which the MSO prefers to improve OTT quality rather than diminish it. The MSO's incentive to improve OTT quality is stronger when it can capture more of the surplus associated with the new content.

In a variety of industries, like healthcare and e-commerce, where consumers access products or services through a gatekeeper platform, the gatekeeper faces a trade-off between steering consumers' choices within the platform towards more profitable products, and allowing consumers free choice among the platform's products with the intention to capture the surplus this generates. Our results do not directly speak to these markets, but our framework provides intuition on the nature of their trade-offs.

Related literature At a high level, our paper relates both to papers that study the market for cable TV (Crawford and Shum, 2007; Crawford and Yurukoglu, 2012; Crawford et al., 2018, 2019), and those that study the market for internet services using high-frequency data (Nevo et al., 2016; Malone et al., 2016, 2014).[3] Our contribution relative to these papers is that we model both TV subscriptions and internet use, the interaction between them, and how the availability of OTT impacts the pricing and quality-provision incentives of MSOs.

Our work is closely related to two other papers that study similar questions but use different data and methods. Malone et al. (2021b) use different data to study consumer behavior on the internet after cutting the cord, namely dropping TV service. The motivation is similar, but their analysis is purely descriptive and they do not estimate consumer preferences or conduct counterfactual analysis as we do here. McManus et al. (2023) use additional data and a difference-in-differences design to measure the effect of UBP's introduction on households' subscription and usage decisions, which complements the estimation and counterfactual analysis in this paper.

Relationships between MSOs and internet content providers are an active area for public policy, especially concerning merger approval and net neutrality.[4] These policy issues converge in vertical mergers between MSOs and media companies, which can affect MSOs' profits from various content sources and therefore induce steering activity. The literature on these issues largely began with Wu (2003), who introduced the term "net neutrality" and provides a summary of the issues. Lee and Wu (2009) and Greenstein et al. (2016) discuss and review the literature on the topic. However, most of the existing economic analysis of the topic is theoretical.[5] Our empirical analysis on steering incentives complements these theoretical studies by providing insight into relevant trade-offs for the debate. Goetz (2019) and Tudon (2021) also make recent related empirical contributions. Goetz (2019) studies the how bargaining between internet service providers and Netflix affects mergers. Tudon (2021) examines the trade-off between content providers' entry and congestion on Amazon's Twitch.

[3]Other studies of demand for broadband services include Prince and Greenstein (2017), Goolsbee and Klenow (2006), Dutz et al. (2009), Rosston et al. (2013), Greenstein and McDevitt (2011), Edell and Varaiya (2002), Varian (2002), and Hitte and Tambe (2007).

[4]The FCC's 2015 Open Internet Order prevented MSOs from discriminating among various online applications. This order limited MSOs' ability to reduce usage of video services from some third-party providers. The FCC voted in 2017 to roll back the order, and future policy in this area continues to be debated.

[5]Economides and Hermalin (2012); Armstrong (2006); Bourreau et al. (2015); Choi et al. (2015); Choi and Kim (2010); Economides and Tag (2012); Gans (2015); Economides and Tag (2016); Reggiani and Valletti (2016); Sidak (2006).

Prior work has examined gatekeeper firms' strategic efforts to steer and sort heterogeneous consumers across product menus in other industries. Ho and Lee (2019), Liebman (2017), and Raval and Rosenbaum (2017) study how insurers influence patients' choices across medical providers. Barwick et al. (2017) examine conflicts of interest and steering by residential real-estate brokers. Crawford et al. (2018) consider similar incentives in cable TV markets and estimate the value to cable distributors of including vertically integrated versus non-integrated sports networks in their channel bundles. Lee and Musolff (2021) examine Amazon's influence on market structure and welfare, as it seeks to balance sales of its own goods against entry of sellers that increase the platform's attractiveness to consumers. Raval (2022) studies Amazon's ability to steer consumers to its own products and services through the Buy Box default purchase option.

The incentive to degrade product quality for discriminatory or steering purposes, as is present in our model, is related to the classic work of Mussa and Rosen (1978), which Crawford and Shum (2007) apply in the context of the telecommunications industry. Contrary to the incentive to degrade, Crawford et al. (2019) find quality higher than is socially optimal in the provision of cable TV. In the bundling literature, Armstrong (2013) and Gentzkow (2007) study how the consumption of one product in a bundle affects utility from other products, which is similar to the relationship between OTT and TV that we study. Chu et al. (2011) and Crawford and Yurukoglu (2012) empirically explore how variations on bundling and other pricing strategies can affect firm profit and consumer welfare. Nonlinear pricing strategies similar to those we examine have been studied in broadband markets (Economides and Hermalin, 2015; Lambrecht et al., 2007), phone service contracts (Miravete, 2003; Grubb, 2015; Grubb and Osborne, 2015), and other markets (Hagemann, 2017; McManus, 2007).

2 Model

In this section, we first introduce a consumer choice model that captures the central benefits consumers derive from MSO subscriptions. The model's first stage mirrors the classic two-good mixed bundling model: consumers purchase subscriptions to individual MSO services (internet and TV) or a bundle of both services. In a second stage, consumers make internet and TV consumption choices to maximize utility on the chosen plan, which includes the opportunity to access some video content from online streaming services. Together, the two stages capture both the intensive and extensive margins of consumer decisions that are relevant to MSO pricing. We then use the model to illustrate the trade-offs faced by the MSO in pricing internet and TV services.

2.1 Setup

We consider a market in which an MSO offers consumers a menu of subscription plans, indexed by k. The firm offers a single TV (t) plan, K_i internet (i) access plans, and K_i bundles (b) of the TV plan with an internet plan. Thus, in total the menu consists of $K = 2K_i + 1$ plans. Plan k has subscription price f_k and, if it includes internet service, connection speed s_k. Additionally, each plan may include an internet usage price schedule $\mathcal{P}_k$ if the MSO uses UBP. The usage prices of the full menu of internet plans are collected in $\boldsymbol{\mathcal{P}}$.

We assume that there are two types of content, indexed by j. Internet (or online-only) content is content type 1, and it is only available with an internet subscription. For example, content such as Netflix original programming, which is only available online, is part of type 1 content. TV (or video) content is type 2 and is available with a TV plan or, to some extent, OTT through a subscription to an online service that requires an internet connection. We use q_1 to denote an internet content consumption level, and $q_2 = q_{2,i} + q_{2,t}$ to denote video consumption levels, where $q_{2,i}$ is content received over the internet and $q_{2,t}$ is received via TV. We use $\mathcal{O}(q_1, q_{2,i}; \mathcal{P}_k)$ to denote the total internet usage-related payment associated with consumption levels q_1 and $q_{2,i}$ on plan k. If plan k does not include UBP, usage payments are zero.

2.2 Preferences and Consumer Choice

Consumers make choices in two stages. First, they choose from the MSO's subscription menu or the outside option. The subscription choice is made in anticipation of the benefit derived from access to a selected option, k. In the second stage, consumers make usage choices to maximize utility from the services available to them in k. Subscription prices are paid in the first stage and any usage-based overage fees are paid in the second stage. The household solves its full choice problem once per month; we suppress the time subscript for convenience. There is no discounting between stages. Preferences are characterized by a consumer-specific parameter vector $\theta = (v_1, v_2, \delta, \phi, \lambda, \alpha) \in \Theta$, the support of tastes across all consumers. We describe these preference parameters below, as we work backwards through the two-stage consumer choice problem.

2.2.1 Stage 2: Usage Choices

In the second stage, consumers choose q_1 and q_2 for the month to maximize utility. For each content type j, the consumer receives marginal utility of $1/\mu$ per unit of content up to a satiation value. The value of μ is an independent monthly draw from an exponential distribution $F(\mu|\lambda)$, where λ is a consumer-specific parameter. The μ value applies to both internet and video consumption, and can be interpreted as capturing (the inverse of) the consumer's per-unit value of time for content available through the MSO's services.[6] Our timing assumption for μ allows a consumer's usage to vary substantially across months while the consumer rationally remains with a fixed plan k.

The utility satiation values are specific to each consumer and content type. For internet content, the consumer-specific utility satiation value is equal to v_1 and is achieved by consuming a quantity of μv_1. Therefore, the utility from q_1 units of content type 1 is $w_1(q_1) = \min(\frac{q_1}{\mu}, v_1)$. For a TV subscriber, the consumer-specific utility satiation value from video content is equal to v_2 and is achieved by consuming a quantity of μv_2.

To capture the presence of OTT, we assume that consumers can receive some fraction, $\delta \in [0, 1]$, of the TV content they value through OTT services over the internet. Thus, a consumer without a TV subscription has satiation utility of δv_2 from video content that is achieved by consuming a quantity of $\mu\delta v_2$. The parameter δ is consumer-specific and captures a combination of OTT availability and the consumer's net benefit from the available OTT content.[7]

Bringing together these cases for video access, utility from q_2 units of type 2 content depends on the consumer's plan k:

$$w_2(q_2) = \begin{cases} \min(\frac{q_2}{\mu}, v_2), & \text{if } k \text{ includes a TV subscription} \\ \min(\frac{q_2}{\mu}, \delta v_2), & \text{if } k \text{ is an internet-only plan} \end{cases}$$

We assume that a consumer's total gross utility is the sum of content-specific utilities: $w(q_1, q_2) = w_1(q_1) + w_2(q_2)$.

Conditional on selecting plan k and observing the current period's μ, the consumer chooses values of q_1 and q_2 to maximize utility while accounting for any usage-related

[6] A more general model would assign a different parameter value for TV and internet usage. However, since we only observe internet usage, in gigabytes, we cannot identify a separate parameter for TV usage.

[7] This can be viewed as a scenario where a consumer enjoys a number of distinct shows available on TV, that yield a utility of v_2, but only a fraction δ of the shows are available through OTT. We do not model consumers' choices across third-party OTT subscription services. We effectively hold these services' characteristics fixed throughout our analysis, while assuming that consumers do not subscribe to these services when the same content is available to them on TV.

charges in $\mathcal{P}_k$. An internet-only subscriber may consume video content over the internet only, so $q_2 = q_{2,i}$. To simplify the consumption choices of bundle subscribers, we assume that they use their TV subscription to access all content that is available via both internet and TV, so $q_{2,i} = 0$.[8] The optimal usage levels for both plan types satisfy

$$\begin{bmatrix} q_1^*(\mu;\theta,\mathcal{P}_k) \\ q_2^*(\mu;\theta,\mathcal{P}_k) \end{bmatrix} = \underset{q_1,q_2}{\operatorname{argmax}} \Big\{ w(q_1,q_2) - \alpha \times \mathcal{O}(q_1,q_{2,i};\mathcal{P}_k) \Big\}, \tag{1}$$

where α is the marginal utility of income. Our notation for the optimal usage levels makes explicit that they depend on the marginal utility draw μ, consumer-specific parameter vector θ, and price schedule $\mathcal{P}_k$.

In the absence of UBP, the consumer chooses values of q_1 and q_2 exactly equal to their satiation levels.[9] A consumer with internet service chooses $q_1^*(\mu;\theta,\mathcal{P}_k) = \mu v_1$ and receives utility $w_1(q_1^*(\mu;\theta,\mathcal{P}_k)) = v_1$. A consumer with TV service chooses $q_{2,t}^*(\mu;\theta,\mathcal{P}_k) = \mu v_2$ for utility $w_2(q_{2,t}^*(\mu;\theta,\mathcal{P}_k)) = v_2$. If the consumer has internet service but not TV, they choose $q_{2,i}^*(\mu;\theta,\mathcal{P}_k) = \delta\mu v_2$ for utility $w_2(q_{2,i}^*(\mu;\theta,\mathcal{P}_k)) = \delta v_2$. Consumption quantities vary with the μ realization but utility does not.

When the MSO uses UBP, the consumer takes into account the marginal price of usage, given by the price schedule $\mathcal{P}_k$, and the marginal benefit of additional usage, determined by the realization of μ, when choosing usage levels. With a positive usage price this may generate $q_1^*(\mu;\theta,\mathcal{P}_k) < \mu v_1$ or $q_{2,i}^*(\mu;\theta,\mathcal{P}_k) < \mu\delta v_2$. Equal marginal utility of usage across content types implies that there may be a continuum of optimal q_1 and $q_{2,i}$ choices when these values are below the consumer's satiation values, but the total optimal internet usage level is unique. We resolve this issue by assuming that the consumer divides internet usage in proportion to the satiation values for each content type, e.g. q_1 is the fraction $[v_1/(v_1 + \delta v_2)]$ of optimal total usage.

Two useful objects for what follows are the expected utility and expected usage-related charges from optimal usage levels. Prior to the realization of μ, the consumer knows λ and the optimal usage choices and associated utility from each potential draw of μ, but not μ's realized value. To make an optimal subscription choice, the consumer must calculate the expected gross utility from usage,

$$w^*(\theta,\mathcal{P}_k) \equiv E_\mu[w(q_1^*(\mu;\theta,\mathcal{P}_k), q_2^*(\mu;\theta,\mathcal{P}_k)].$$

[8] When internet usage is costly, this assumption represents the best-case/least-cost outcome for the MSO and strengthens incentives to steer consumers to the bundle.

[9] This assumption guarantees a unique solution to the utility maximization problem and could instead be rationalized explicitly with a vanishingly-small opportunity cost of time.

Similarly, the expected usage-related payment is

$$\mathcal{O}^*(\theta, \mathcal{P}_k) \equiv E_\mu[\mathcal{O}(q_1^*(\mu; \theta, \mathcal{P}_k), q_{2,i}^*(\mu; \theta, \mathcal{P}_k); \mathcal{P}_k)].$$

Our assumptions on a consumer's tastes and usage choices imply a distribution of content-specific and total usage choices conditional on θ. Let $G_j(q; \theta, \mathcal{P}_k)$ represent the ex-ante probability that a consumer with tastes θ who selects subscription k chooses an optimal usage level $q_j^*(\mu; \theta, \mathcal{P}_k)$ which is less than q. This probability distribution takes its value through optimal usage choices for potential realizations of μ:

$$G_j(q; \theta, \mathcal{P}_k) = \int_0^\infty \mathbb{1}[q_j^*(\mu; \theta, \mathcal{P}_k) < q] dF(\mu; \lambda).$$

We write $g_j(q; \theta, \mathcal{P}_k)$ as the density function that corresponds to $G_j(q; \theta, \mathcal{P}_k)$. The distribution of total internet usage is defined similarly:

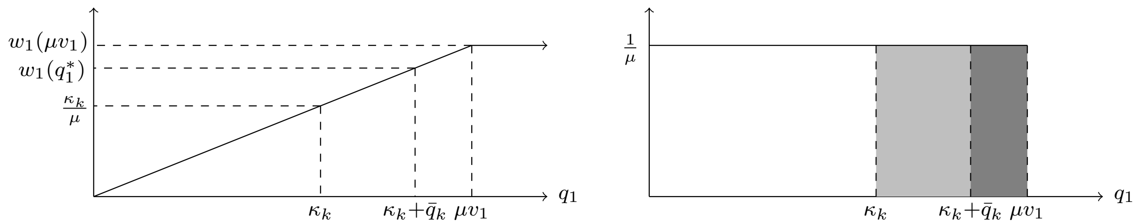
$$G(q; \theta, \mathcal{P}_k) = \int_0^\infty \mathbb{1}[q_1^*(\mu; \theta, \mathcal{P}_k) + q_{2,i}^*(\mu; \theta, \mathcal{P}_k) < q] dF(\mu; \lambda), \tag{2}$$

with corresponding density $g(q; \theta, \mathcal{P}_k)$. The densities may include mass points and jumps due to nonlinear aspects of the price schedule $\mathcal{P}_k$ even though μ is drawn from an exponential distribution.

Before moving on to discuss the consumer's subscription choice, we describe the price schedule in our data and illustrate how consumers' preferences interact with it to generate usage choices. We provide a simple case here and the full details in Appendix A. The empirical price schedule, $\mathcal{P}_k^e$, consists of a plan-specific usage allowance κ_k and a "top-up" fee $\overline{p}_k$ for each discrete increase of the allowance of size $\overline{q}_k$ gigabytes (GBs). Under $\mathcal{P}_k^e$, total internet usage $q = q_1 + q_{2,i}$ generates UBP expenditure of $\mathcal{O}(q_1, q_{2,i}; \mathcal{P}_k^e) = 0$ if $q \leq \kappa_k$ and $\mathcal{O}(q_1, q_{2,i}; \mathcal{P}_k^e) = n\overline{p}_k$ when usage of q requires the purchase of n top-ups beyond κ_k.

Consider the optimal choices of a consumer who subscribes to an internet-TV bundle, and therefore the only internet usage is $q_1^*(\mu; \theta, \mathcal{P}_k)$. In Figure 1 we illustrate the optimal usage decision of such a household, and the corresponding values of $w_1(q_1^*(\mu; \theta, \mathcal{P}_k))$ and $\mathcal{O}(q_1^*(\mu; \theta, \mathcal{P}_k), 0; \mathcal{P}_k^e)$. In panel (a) of Figure 1, utility increases with q_1 until it reaches the satiation level, μv_1. We indicate on the horizontal axis the plan allowance κ_k and the allowance plus one top-up (i.e., $\kappa_k + \bar{q}_k$). The household's decision is whether to use just the allowance, or purchase one or more top-ups to the allowance, each with price $\bar{p}_k$. The household only realizes the full v_1 if it chooses the satiation level of usage (i.e., where utility plateaus). The optimal level of usage is easiest to see in panel (b) of Figure

Figure 1: Utility from Usage



(a) Utility

(b) Marginal Utility

Notes: Utility (left) and marginal utility (right) from internet consumption for a household that chooses the bundle.

1, which shows the marginal utility of usage, which equals $1/\mu$ up to μv_1 and zero after. Purchasing one top-up is optimal if

$$\frac{\bar{q}}{\mu} \geq \alpha \bar{p}_k \quad \text{and} \quad \frac{\mu v_1 - \kappa_k - \bar{q}}{\mu} < \alpha \bar{p}_k.$$

In panel (b), these conditions hold when the lighter shaded region is greater than $\alpha \bar{p}_k$, and the darker region is smaller than $\alpha \bar{p}_k$. For this realization of μ, $w_1(q_1^*(\mu;\theta,\mathcal{P}_k)) = \frac{\kappa_k+\bar{q}}{\mu}$ and $\mathcal{O}(q_1^*(\mu;\theta,\mathcal{P}_k),0;\mathcal{P}_k^e) = \bar{p}_k$.

2.2.2 Stage 1: Subscription Choices

At the start of each period, prior to the usage stage, consumers make a subscription decision. A household's expected utility u_k from MSO menu option k is the sum of its expected utility from usage, $w^*(\theta,\mathcal{P}_k)$; its benefit from greater internet connection speed, ϕ/s_k; its disutility from subscription price and expected overage charges, $\alpha\Big(f_k + \mathcal{O}^*(\theta,\mathcal{P}_k)\Big)$; and an idiosyncratic taste shock associated with plan k, ϵ_k:

$$u_k = w^*(\theta,\mathcal{P}_k) - \phi/s_k - \alpha\Big(f_k + \mathcal{O}^*(\theta,\mathcal{P}_k)\Big) + \epsilon_k. \tag{3}$$

The functional form ϕ/s_k implies that a consumer's benefit from speed is increasing in s_k at a decreasing rate. The marginal disutility of payments to the MSO, α, applies to both the fixed fee and expected overage charges.

If the consumer selects the outside option, they receive the $u_0 = \epsilon_0$. The values of ϵ associated the MSO's menu and the outside option are distributed i.i.d type-I extreme value.

To make their subscription decision, the consumer evaluates the utility associated with each MSO menu item and the outside option, and they select the option that provides the greatest value. The distributional assumption on the ϵ_k values and ϵ_0 imply that the consumer's probability of subscribing to plan k is:

$$\Pr(d_k = 1|\theta, \boldsymbol{\mathcal{P}}) = \frac{\exp\left(w^*(\theta, \mathcal{P}_k) - \phi/s_k - \alpha\Big(f_k + \mathcal{O}^*(\theta, \mathcal{P}_k)\Big)\right)}{1 + \sum_m \exp\left(w^*(\theta, \mathcal{P}_m) - \phi/s_m - \alpha\Big(f_m + \mathcal{O}^*(\theta, \mathcal{P}_m)\Big)\right)}, \tag{4}$$

where $d_k = 1$ when a consumer subscribes to plan k and $d_k = 0$ otherwise, with $\sum_k d_k = 1$. Equation (4) and the density of total internet usage, $g(q; \theta, \mathcal{P}_k)$ form the basis for our estimation approach, which seeks to match subscription and usage choices of households to values of θ that are most likely to have generated such behavior.

2.3 Pricing Incentives

An MSO may use UBP in pursuit of two goals that are not fully attainable with lump-sum subscription fees, f_k. First, to the extent that consumer surplus increases in internet usage (including OTT), the MSO has an incentive to collect additional revenue by metering usage through UBP. Second, the firm may respond to differences in internet and TV costs by steering consumers through usage-based prices. While subscription prices steer consumers among services as well, steering through UBP can be complementary and may be effective on additional margins (e.g., usage volume, consumer-specific costs of service).

To explore the distinct metering and steering incentives in UBP, we specify a usage price schedule with per-unit usage prices which capture the MSO's key trade-offs. We consider a novel UBP schedule, $\mathcal{P}_k^{\tau}$, that includes an all-purpose per-unit price, τ, that applies to all internet content, and an OTT-specific price, τ_δ, which can raise or lower the price of video delivered via internet. UBP expenditure for internet usage $q_1 + q_{2,i}$ is therefore $\mathcal{O} = \tau q_1 + (\tau + \tau_\delta) q_{2,i}$. We focus on usage prices that are constant across internet plans, captured by k.

To build intuition, consider an MSO that offers a single internet service plan and has three subscription options: internet, TV, and a bundle of the two. The MSO faces lump-sum per-consumer internet and TV subscription costs c_i and c_t, respectively; the cost of providing the bundle is $c_b = c_i + c_t$. The firm's subscription prices f_i, f_t, and f_b lead to choices by consumers which generate market shares equal to S_i, S_t, and S_b. Let Q_i represent the sum of all consumers' total internet usage, which is $q_1 + q_{2,i}$ per consumer.

Total OTT usage is $Q_{2,i}$, which is the sum of all $q_{2,i}$. Normalizing the consumer population to one, the firm's profit is:

$$\pi = (f_i - c_i)S_i + (f_t - c_t)S_t + (f_b - c_i - c_t)S_b + \tau Q_i + \tau_\delta Q_{2,i}.$$

We can illustrate consumers' choices as a simple extension of the classic mixed bundling model if we assume that consumers differ only in their tastes $v = (v_1, v_2)$, which are distributed on $[0,1] \times [0,1]$. Specifically, we fix $\mu = 1$, eliminate variation due to ϵ_k, and assume that consumers have common preferences for OTT and speed, $\delta > 0$ and $\phi > 0$, respectively. To begin, we set $\tau = \tau_\delta = 0$, so the MSO receives revenue from subscription payments only. In Figure 2 we sketch the separation of consumers by v for subscription prices with the following relationships: $f_b > f_i$, $f_b > f_t$, and $f_b < f_i + f_t$. Consumers with v in area i choose an internet-only subscription, those in t chose TV only, and those in b select the bundle. One impact of δ is seen in margin II, which has slope $(1-\delta)^{-1}$. Holding prices fixed, an increase in δ moves some consumers from 0 to i. Another impact is seen along margin III, the location of which shifts up with an increase in δ as some consumers move from b to i and consume video via improved OTT rather than TV.[10]

Without usage fees, all consumers in a single subscription region pay the same price but have different usage and surplus. Consider the consumers who select i. Non-video internet usage and surplus increase with v_1 (moving horizontally), while OTT usage and surplus increase in v_2 (moving vertically). Now consider the firm's incentive to increase τ from zero, therefore adding a price for all content delivered over the internet. Given our assumptions on usage-related utility (i.e., $\mu = 1$), a small τ does not affect any subscribers' usage choices or have a first-order impact on subscription decisions, but it reallocates surplus to the MSO by acting on all the inframarginal consumers in regions i and b in Figure 2. These consumers' willingness-to-pay generate the MSO's incentive to meter usage through τ. Each consumer in i pays $\tau(q_1 + q_{2,i})$, and each consumer in b pays τq_1.

As τ increases away from 0, holding subscription fees fixed, there are meaningful impacts on all margins where positive internet usage affects consumer choices.[11] The MSO loses some consumers to the outside option (region 0) from the bundle (margin I) and internet-only (margin II). Some consumers also move from i to b (margin III) to reduce their internet usage via OTT and therefore their usage-related expenditure; some b subscribers switch to t when their internet usage becomes more expensive (margin IV).

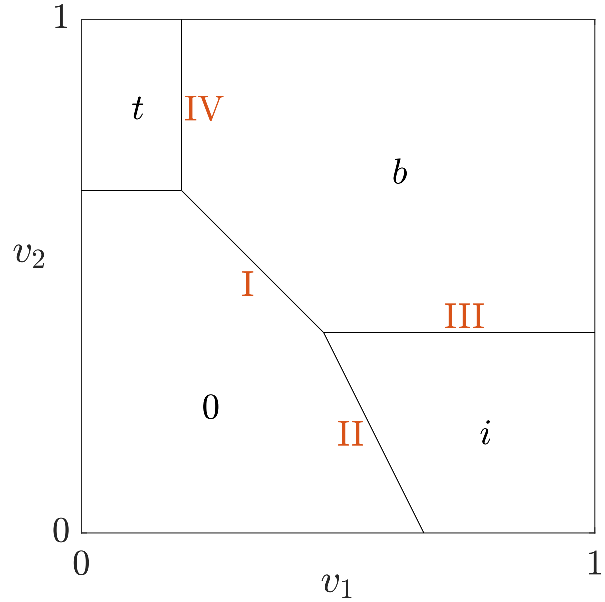
[10] A consumer on margin III has taste values which satisfy $(1-\delta)v_2 = f_b - f_i$.

[11] We hold fixed the subscription prices in this discussion to make the intuition simpler.

When a consumer with OTT usage $q_{2,i} > 0$ moves from i to b, this changes MSO profit by $(f_b - c_b) - (f_i - c_i) - \tau q_{2,i}$. The difference in bundle versus internet-only subscription profit margins impacts how the MSO views the cost of lost usage revenue, $\tau q_{2,i}$. This leads naturally to the potential benefits of an OTT-specific price, τ_δ, so that the per-unit price for OTT is $(\tau + \tau_\delta)$. When $\tau_\delta > 0$, the additional price extracts more surplus from OTT usage by i subscribers while accelerating movement across margin III because of greater savings for consumers who use $q_{2,i} > 0$. This additional incentive for consumers to shift to b is more beneficial to the MSO for greater values of $(f_b - c_b) - (f_i - c_i)$. On the other hand, if TV is relatively unprofitable for the MSO, the firm can use $\tau_\delta < 0$ to steer consumers back toward i despite $\tau > 0$. Relative to reducing τ, a value $\tau_\delta < 0$ entails different trade-offs for the MSO because consumers at the margin between i and b have greater taste for video content than inframarginal i subscribers. In terms of the simple model illustrated in Figure 2, consumers closer to the horizontal axis use less OTT, so losses on OTT usage revenue from this group are limited. This is a novel channel in the discussion of UBP: conditional on satisfying the general metering incentive that drives τ, an MSO may use additional (negative) UBP to provide advantages for OTT relative to other internet usage.

Similar intuition applies to an MSO's incentive to alter δ. Relative to a firm that sells only internet subscriptions, an MSO that offers b bears less cost from a reduction in δ because some consumers who drop their i subscriptions will move to the MSO's own b instead of the outside option. When the relative profit margins of the bundle versus the internet are such that the MSO has an incentive to steer toward the bundle (i.e., $f_b - c_b > f_i - c_i$), a reduction in δ contributes to moving consumers across Figure 2's margin III. With usage prices, however, the profitability of internet subscriptions increases, and there is less incentive to reduce δ. Usage prices, in fact, may reward an MSO from increasing δ because of their positive impact on the profitability of internet subscriptions. In the $\mathcal{P}_k^\tau$ we consider, an increase in δ provides the firm with a return proportional to $(\tau + \tau_\delta)$ for each internet subscriber.

Figure 2: Consumer choices in simplified model



3 Data

Our data come from one MSO and describe activity for approximately 9 months in a large North American city. During our sample, the MSO is the sole provider of high-speed internet service for almost all households in the market; alternatives include low-speed DSL internet service and satellite TV service. We do not observe subscriptions and usage for households that choose these alternatives. We observe 34,752 households' billing information, subscriptions, and internet usage.[12]

Like most North American MSOs, the firm we observe offers internet and TV services via mixed bundling, giving discounts off standalone prices when consumers subscribe to both services.[13] Across internet tiers, the average price difference between the bundle and internet-only subscriptions is about \$100. 23% of the MSO customers have internet-only subscriptions, and the remaining 77% subscribe to an internet-TV bundle; no MSO customers in our data subscribe to TV alone. The MSO offers tiers of internet service differentiated by speed and, as we discuss below, usage allowance during part of the sample period. The typical price difference between adjacent internet tiers is about \$15, and is the same with or without a TV subscription. In the first row of Table 1 we present the

[12]Our agreement with the MSO prevents us from identifying the firm or any details that could be used to infer its identity. This includes the specific market served, the exact dates and details of the implementation of UBP, as well as detailed characteristics of the MSO's product offerings.

[13]The MSO also offers telephone service, which about 40% of its customers subscribe to. We do not use the telephone service information in this paper.

share of households who choose different plans, aggregated by plan speed.

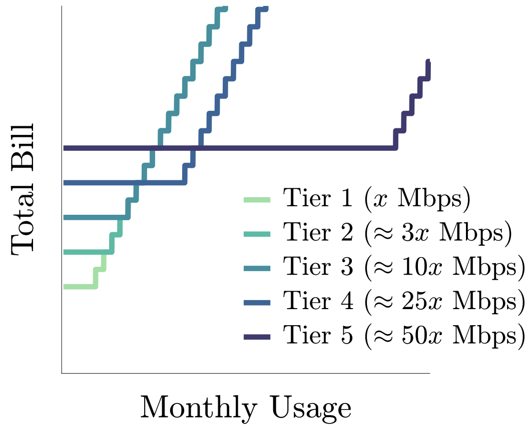
For each household in the sample, we observe download and upload volume each month, which we aggregate into total household monthly usage (in gigabytes). In Table 1 we show descriptive statistics on monthly usage by subscription type. The mean (median) monthly usage level across all households in the sample is 105 (49) gigabytes, or about 3.5 (1.6) gigabytes per day. Internet usage differs substantially across households. Average usage in the highest-priced (and highest-speed) tier is nearly seven times that in the lowest-priced tier. Within-tier usage dispersion is also substantial; coefficients of variation range from 1.67 to 2.05 across tiers. Across combinations of TV and internet service, internet-only subscribers have mean (median) internet usage 61% (137%) greater than bundle subscribers. There is also substantial variation in usage across months within a household. Decomposing the variance in usage across all subscriber-months, 83% of variation is explained by heterogeneity between households, while the remaining 17% is explained by within-household variation.

UBP was introduced in this market in the middle of our sample period. To our knowledge, the market was chosen for the introduction of UBP due to the network characteristics that permitted billing on usage. The timing of UBP's introduction was based on engineering considerations, not local demand conditions that would impact the return from UBP.[14] The MSO implemented UBP by attaching a multipart price schedule to each of its existing service tiers, holding all other characteristics (i.e., subscription fees and connection speeds) fixed. Tiers vary in their monthly usage allowance in GBs. Usage up to this allowance is included in the monthly subscription charge, but if a household exceeds its allowance, it is charged for an extra top-up of data, which the consumer may use fully or partially. In Figure 3 we illustrate the total price consumers pay for different monthly usage depending on their plan.

The MSO's introduction of UBP came in two phases. The MSO announced that it would implement UBP starting on a specified date, and it provided households with information about how their monthly usage compared to the data allowance of their current internet tier. During this announcement phase, households were not billed if their usage exceeded their tier's allowance. In the next phase, which we call the "UBP period," the MSO assessed overage charges on households that exceeded their allowances. We observe several months of activity (a "pre-UBP period"), immediately prior to the announcement phase. Each period spans multiple monthly billing cycles. For the analysis below, we use

[14] Competitors' subscription offerings did not change meaningfully during the sample period, in response to the UBP policy's introduction. Satellite TV was available in the market, as was a substantially slower alternative internet service via DSL.

Figure 3: Cost of Usage by Tier Under UBP



Notes: We do not provide numerical labels on the axes to preserve the MSO's anonymity.

data from the pre-UBP period and the UBP period, but not the announcement period.

The third panel of Table 1 shows that some internet-only consumers added TV service after the introduction of UBP. However, upgrades of internet service tiers were much more common, especially for consumers who had already selected high speed tiers during the pre-UBP period. The bottom panel of Table 1 shows that 3% of household-month usage levels during the pre-UBP period would have resulted in overage charges after the price change, with an average bill of $32 conditional on exceeding the usage allowance. Consistent with the expectation of overage charges, 9% of households upgraded their initial internet tier to a tier with a higher usage allowance.

Table 1: Descriptive Statistics

	Internet-only			Internet & TV		
Speed Tier	Low	Median	High	Low	Median	High
Choice Share	0.07	0.12	0.04	0.18	0.50	0.09
Monthly Usage						
Mean	98.90	171.23	307.56	43.18	84.54	171.20
Standard Deviation	106.48	151.90	266.07	71.46	116.83	208.23
5th Percentile	2.86	13.55	30.95	0.42	1.49	3.98
25th Percentile	22.91	62.53	119.16	4.38	12.09	28.77
Median	65.83	132.87	244.07	15.04	39.18	94.94
75th Percentile	138.78	236.95	424.70	49.52	112.63	247.01
95th Percentile	305.11	452.08	759.40	185.88	313.12	569.83
Subscription Changes						
Upgrade Tier	0.06	0.08	0.21	0.07	0.06	0.22
Add Video	0.03	0.03	0.03	–	–	–
Price Change Impact						
Share w/ Overages	0.04	0.08	0.05	0.02	0.02	0.01
Mean Overage ($)	23.44	31.93	33.60	24.88	34.03	43.69
Median Overage ($)	20.00	20.00	30.00	20.00	20.00	30.00
Observations	22,773	36,994	12,550	57,026	156,171	27,252

Notes: Summary statistics at a subscriber-month level of observation using 312,678 observations across 34,752 households and 9 months. The first two panels contain shares and usage levels (measured in gigabytes) by subscription type. The third panel contains the fraction of households who changed their pre-UBP period subscriptions during the announcement or UBP periods. The final panel describes overage charges that would have resulted from applying the UBP period billing schedule to pre-UBP period usage levels, with the means and medians conditional on positive overage charges.

4 Econometrics

In this section, we outline our approach to estimating the model presented in Section 2 using the subscription and usage data described in Section 3. Our goal is to recover the distribution of preferences across households, where each household is described by a vector of six parameters, $\theta = (v_1, v_2, \phi, \lambda, \delta, \alpha)$. Given the substantial heterogeneity in observed decisions, we also seek to limit assumptions on the distribution of the parameters across households. We therefore use a fixed-grid estimation approach (Ackerberg, 2009; Fox et al., 2011, 2016), which most closely follows the likelihood-based approach of Malone et al. (2021a), to estimate the distribution of preferences across households. In this section we discuss identification, but we first present the estimation strategy, as it helps shed light on how we use the data to recover the distribution of θ.

4.1 Estimation and Inference

We implement a two-step estimation routine. In the first step, we solve the model for a large set of candidate θ types. In the second step, we recover the distribution of θ by assigning probability mass to the types whose predicted behaviors rationalize the observed household decisions.

We begin the first step by drawing $R = 1,000,000$ candidate types from the type space Θ, with r indexing individual values. Specifically, we fix the support of Θ after experimenting to ensure that the full range of observed behaviors is rationalized. We then draw R types from the fixed support using the six-dimensional Halton sequence. This gives us a set, $\theta_r = (v_{1r}, v_{2r}, \phi_r, \lambda_r, \delta_r, \alpha_r)$, $r = 1 \ldots R$, of different types.

We solve the model for each type r. The model solution provides type r's subscription probabilities for each option k, as in equation (4), and a density of total internet usage conditional on k, which is based on the distribution in equation (2). To calculate these two objects, we use simulation to account for the distribution of μ conditional on λ_r. For each type r, we draw 50,000 μ values from $F(\mu|\lambda_r)$ and solve for optimal usage quantities, $q_j^*(\mu; \theta_r, \mathcal{P}_k^e)$, for each observed subscription option k with corresponding UBP schedule $\mathcal{P}_k^e$. We use the collection of usage solutions to simulate the density $g(q; \theta_r, \mathcal{P}_k^e)$. The probability of any realized usage level is a function of parameters v_{1r}, v_{2r}, δ_r, λ_r, and, under UBP, α_r. In the absence of UBP, this dependence is simple: optimal usage given a particular realization of μ is exactly equal to $\mu(v_{1r} + \delta v_{2r})$ for an internet subscriber and μv_{1r} for a bundle subscriber. With UBP, α_r enters as well, resolving trade-offs between the marginal utility of usage in excess of a plan's allowance its marginal price, determined by $\mathcal{P}_k^e$.

In addition to using this density to predict usage quantities in the estimation approach's second step, we use it to calculate the expected values $w^*(\theta_r, \mathcal{P}_k^e)$ and $\mathcal{O}^*(\theta_r, \mathcal{P}_k^e)$, where we add r subscripts to highlight the dependence on θ_r. When optimal usage is not truncated by UBP, $w^*(\theta_r, \mathcal{P}_k^e)$ equals $v_{1r} + \delta_r v_{2r}$ for internet-only plans and $v_{1r} + v_{2r}$ for bundle plans. Once we have obtained $w^*(\theta_r, \mathcal{P}_k^e)$ and $\mathcal{O}^*(\theta_r, \mathcal{P}_k^e)$, it is straightforward to apply equation (4) and calculate $\Pr(d_k = 1; \theta_r, \boldsymbol{\mathcal{P}}^e)$, the probability that a type r consumer selects subscription option k.

After solving the model for each of the R candidate types, we store the subscription choice probabilities and corresponding optimal usage densities for both price schedules we observe in our data. We then proceed to the second step of the estimation procedure. In this step, we compute the household-specific likelihood that a household's realized subscription and usage choices were generated by a specific type's parameters. By starting with a discrete set of types, we are implicitly assume a prior distribution with uniform probability mass on each discrete types we draw and zero mass elsewhere. We then apply Bayes' rule to update the initial prior on the probability distribution of types describing each household, then aggregate across households to obtain a population-level posterior distribution of θ.

For each household $h = 1, \ldots, N$ in the sample, we observe a sequence $m = 1, \ldots, M$ of monthly subscription and usage decisions. When household h selects subscription k in month m, we set $d_{khm} = 1$. We observe q_{khm} as household h's observed total internet usage during m while in subscription k. The likelihood that an observed sequence of household decisions was generated by a candidate type with preference parameters θ_r is

$$\mathcal{L}_h(\theta_r) = \prod_{m=1}^{M} \Pr(d_{khm} = 1; \theta_r, \boldsymbol{\mathcal{P}}_m^e) \times g(q_{khm}; \theta_r, \mathcal{P}_{km}^e)$$

where $\boldsymbol{\mathcal{P}}_m^e$ is the the set of plan-specific price schedules, $\mathcal{P}_{km}^e$, available during month m.

Assuming a uniform prior across the set of candidate types for each household, the relative likelihoods for each household correspond to weights in a discrete posterior distribution. Specifically, the probability that household h is of type r equals

$$\omega_{hr} = \frac{\mathcal{L}_h(\theta_r)}{\sum_{l=1}^{R} \mathcal{L}_h(\theta_l)}.$$

To obtain an estimate of the distribution of types across the population, we aggregate the household-specific posterior weights across households to obtain type weights $\omega_r = \frac{1}{N}\sum_{h=1}^{N}\omega_{hr}$, for $r = 1,\ldots,R$.

Some regions of the sampled type space Θ yield optimal decisions that are inconsistent with households' observed usage (i.e., $\theta_r \in \Theta$ $s.t.$ $\mathcal{L}_h(\theta_r) = 0$ $\forall h$). To ease the computational burden of estimation with no impact to estimates of the type-specific posterior weights (i.e., ω_r for $r = 1,\ldots,R$), we identify those candidate types with $\mathcal{L}_h(\theta_r) = 0$ $\forall h$. We do so by calculating a type's expected pre-UBP usage, which is simple to calculate for all R types, and comparing it to the empirical usage density. All types with expected usage in excess of 4 terabytes per month have $\mathcal{L}_h(\theta_r) = 0$ $\forall h$. This reduces computation to solving the model fully for the remaining 67,753 candidate types.

To calculate standard errors for the type weights ω_r, and associated statistics of the weights, we use block re-sampling at the household level. Specifically, we re-sample the population of households with replacement 500 times and re-compute each candidate type's weight (or function of those weights). This approach is computationally light because it does not require re-solving the model or re-computing household likelihoods. Instead, we simply re-weight each household's contribution to the calculation of each type weight ω_r.

4.2 Identification

We now turn to the topic of identification. Consider a time series of monthly subscription and usage decisions for household h, i.e., $\{d_{khm}, q_{khm}\}_{m=1}^{M}$, with variation in plan features. Assume that household h, which selects subscription k in month m, faces no UBP during the sample. The household's probability of choosing plan k during month m is given by equation (4), and therefore co-variation between subscription decisions and plan features (i.e., price and speed) identify the price coefficient (α), preference for speed (ϕ), and two subscription-type constants, γ_i and γ_b. Without UBP, expected usage varies only between plan types but not with other plan characteristics. Therefore, γ_i is equal to household h's expected utility $w^*(\theta, \mathcal{P}_k)$ for any internet-only plan k, and γ_b is equal to the expected utility of any bundle plan. This implies $\gamma_i = v_1 + \delta v_2$ and $\gamma_b = v_1 + v_2$. Thus, subscription decisions identify α, ϕ, and the quantities $v_1 + \delta v_2$ and $v_1 + v_2$.

The time series of internet usage decisions, q_{khm}, conditional on subscription decisions, d_{khm}, identifies λ and separately identifies v_1, v_2, and δ. Specifically, when plan prices and speeds are such that the household chooses an internet-only plan, the household selects $q_{khm} = \mu_m(v_1 + \delta v_2)$, where μ_m is a realization of the marginal utility shock drawn from an exponential distribution with mean λ. As noted, $\gamma_i = v_1 + \delta v_2$ is identified from subscription decisions; therefore $\mu_m = \frac{q_{khm}}{\gamma_i}$, and λ is identified as the mean of the sequence of $\frac{q_{khm}}{\gamma_i}$ values. Given λ, usage differences across internet and bundle plans separately identify v_1, v_2, and δ. In particular, the difference in mean usage between internet-only and bundle plans is $\gamma_q = \frac{\delta v_2}{\lambda}$. With γ_i, γ_q, and λ, v_1 is identified because $v_1 = \gamma_i - \lambda\gamma_q$. Similarly, v_2 is identified because $v_2 = \gamma_b - \gamma_i + \lambda\gamma_q$. Finally, δ is identified because $\delta = \lambda\gamma_q/(\gamma_b - \gamma_i + \lambda\gamma_q)$.

In summary, time series variation in household choices as plan attributes change identifies the parameters of the model. In practice, however, one can ask whether the variation we observe in the data is sufficient to pin down the parameters. In the fixed-grid estimation approach this is equivalent to asking: How informative are the observed choices in updating the prior assumption of a uniform distribution of θ? Specifically, the model provides predictions regarding plan choices and usage for each candidate type, and our estimation approach identifies those types that are most likely to have generated each household's sequence of choices. In our sample, the implementation of UBP generates the identifying variation in plan-specific prices and expected utility that allows us to refine the distribution of θ.

We obtain some identifying power from a static choice of internet tier prior to UBP. In this setting there are no usage allowances, so the only reason to choose a more expensive plan is to receive greater speed. A household's tier selection, therefore, is informative about its preference for speed (ϕ) and disutility from price (α). To build intuition, assume there are no ϵ_{rk} terms, i.e., the "logit" shocks to plan-specific utility shown in equation (3). With this abstraction, plan choice is deterministic. Any household that chooses plan k must come from a type that satisfies $u_k^+ < u_k$ and $u_k^- < u_k$, where $^-$ and $^+$ denote "adjacent" plans in terms of speed. Given the linearity of utility in speed and price, this implies

$$\frac{p_k - p_k^-}{1/s_k^- - 1/s_k} < \frac{\phi_r}{\alpha_r} < \frac{p_k^+ - p_k}{1/s_k - 1/s_k^+}. \tag{5}$$

In other words, for a household that chooses plan k, we update our belief about the (joint) values of ϕ_r and α_r that would select plan k. While our prior assumption puts uniform weight on all R types, the selection of k in this simplified example directs us to place zero weight on θ_r with combinations of ϕ_r and α_r that violate the inequalities in (5) and uniform weight on parameter values that satisfy (5). UBP's introduction adds price variation that helps separately identify α_r, but changes in $\mathcal{O}^*(\theta_r, \mathcal{P}_k^e)$ are accompanied by changes to $w^*(\theta_r, \mathcal{P}_k^e)$, which complicate the simple inequalities in (5).

Optimal usage within a plan, the associated $w^*(\theta_r, \mathcal{P}_k^e)$ and $\mathcal{O}^*(\theta_r, \mathcal{P}_k^e)$, and therefore plan choices are determined jointly by v_1, v_2, λ, δ, and α, so identification of these parameters is more nuanced. To build intuition about identification, first consider a household's subscription decision prior to UBP. If a household chooses internet-only, the plan choice reveals information about the quantity $v_1+\delta v_2$. Specifically, the utility derived from unrestricted usage must be sufficient such that the household prefers it to the outside option given speed and price. Similarly, if a household chose a bundle option, the plan choice reveals information about the quantity $v_1 + v_2$.

A household's pre-UBP choice between internet-only and bundled provides further information about the structural parameters determining the payoffs (i.e., v_1, v_2, and δ). Abstracting away from the logit error as above, a household's preference for an internet-only plan over the TV-internet bundle with the same speed implies that $v_1 + \delta v_2 - \alpha f_i > v_1 + v_2 - \alpha f_b$. This is equivalent to $\frac{v_2(1-\delta)}{\alpha} < (f_b - f_i)$, so the observed price difference together with the subscription choice place a bound on $\frac{v_2(1-\delta)}{\alpha}$. In a probabilistic sense, i.e., via the logistic error, persistence or changes in a household's subscription choices provides additional information on the value of these quantities (sums and products of structural parameters) relative to observed price differences.

A household's usage decisions provide further identifying information relevant to the structural parameters that determine utility from usage. Households that choose internet-only and bundle plans generate internet usage levels of $\mu(v_1 + \delta v_2)$ and μv_1, respectively, for a given realization of μ. Conditional on subscription choices identifying one or both of $(v_1 + \delta v_2)$ and $(v_1 + v_2)$, the mean and variance of the household's usage across months on the same plan identify the underlying mean (i.e., λ) of μ's exponential distribution.

For households that switch subscriptions during the panel, usage levels identify each component of the utility from usage. Specifically, if a household switches from an internet-only plan to a bundle plan, average usage decreases by δv_2. This response, together with information on $v_2(1 - \delta)$ and $v_1 + \delta v_2$ from the subscription choices, helps separate the effects of v_1, v_2, and δ in households' choices.

Even among households that do not alter their subscription choices with UBP's introduction, this choice to not switch helps disentangle v_1, v_2, and δ. In particular, consider a household that chooses an internet-only plan before and after the implementation of UBP, and suppose that $\mu(v_1 + \delta v_2)$ often exceeds the plan's allowance, κ_k. Because the household does not upgrade to a higher-allowance tier or switch to the bundle, it must be the case that the lost surplus associated with reduced usage is smaller in magnitude than the utility impact of paying for a different plan. In addition, foregoing a switch to the bundle suggests that unlimited access to video content is not worth the additional costs of adding a TV subscription. This means the initial usage was likely to be driven by large v_1 rather than large δ and v_2. Conversely, upgrading to a higher internet tier rather than switching to the bundle reveals that δ is large relative to v_2. Among households with similar δv_2 levels, those for whom δ is large will prefer to upgrade their tier at a fairly low cost rather than switch to the bundle (at greater cost) because they can realize a large share of video surplus with an internet-only subscription. In Section 5 we show the extent to which different types of household choices lead to greater refinement (i.e., fewer types with positive weight) across household posteriors.

5 Results

We now turn to the results. First, we describe the estimates of the distribution of model parameters. We show that our flexible approach is able to capture the wide range of behaviors observed in the data. Next, we use the estimates to characterize willingness-to-pay for access to OTT and plan characteristics such as internet connection speed. Finally, we study an MSO's incentives to steer consumers, which can occur through technological means that alter δ directly, or through pricing via UBP schedules that target OTT traffic.

5.1 Model Estimates and Fit

The estimation approach described in Section 4 yields weights that characterize a discrete distribution of the parameters. In the top panel of Table 2, we present the means, standard deviations, medians, and 25^{th} and 75^{th} percentiles for each parameter. The wide range of ϕ captures the heterogeneity in preference for speed, which drives selection of consumers across tiers in the absence of UBP, along with α which determines willingness-to-pay for such features. Similarly, the long tail of λ's distribution helps the model match the substantially higher usage of some households. There is also substantial heterogeneity in v_1, v_2, and δ, which collectively play an important role in determining valuations of services and plan selection. See Appendix Figure A1 for graphical displays of each structural parameter's marginal cumulative distribution function.

In the bottom panel of Table 2, we present the correlations between pairs of parameters. There are some intuitive patterns in these correlations. For example, v_1, v_2, δ, and λ, which collectively determine usage, are positively correlated with ϕ. Thus, higher usage households have a greater preference for speed. This pattern is reflected in the pre-UBP period, when households with greater usage selected higher-speed tiers.

Table 2: Summary Statistics of Type Distribution

Marginals	v_1	v_2	δ	ϕ	λ	α
Mean	88.27	178.73	0.35	950.30	1.04	10.63
	(0.01)	(0.01)	(0.00)	(0.19)	(0.00)	(0.00)
SD	78.09	62.50	0.25	884.96	0.65	31.49
	(0.01)	(0.01)	(0.00)	(0.19)	(0.00)	(0.01)
25th Pct.	12.74	145.13	0.13	331.85	0.61	0.35
	(0.01)	(0.06)	(0.00)	(0.17)	(0.00)	(0.00)
50th Pct.	74.68	198.47	0.31	745.74	0.88	1.10
	(0.06)	(0.02)	(0.00)	(0.13)	(0.00)	(0.00)
75th Pct.	149.15	225.41	0.49	1226.34	1.21	5.74
	(0.04)	(0.02)	(0.00)	(0.37)	(0.00)	(0.01)
Correlations	v_1	v_2	δ	ϕ	λ	α
v_1	1.00	-0.26	0.22	0.30	0.54	-0.34
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
v_2		1.00	-0.20	0.05	0.08	-0.04
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
δ			1.00	0.16	0.28	-0.23
			(0.00)	(0.00)	(0.00)	(0.00)
ϕ				1.00	0.36	-0.11
				(0.00)	(0.00)	(0.00)
λ					1.00	-0.19
					(0.00)	(0.00)
α						1.00
						(0.00)

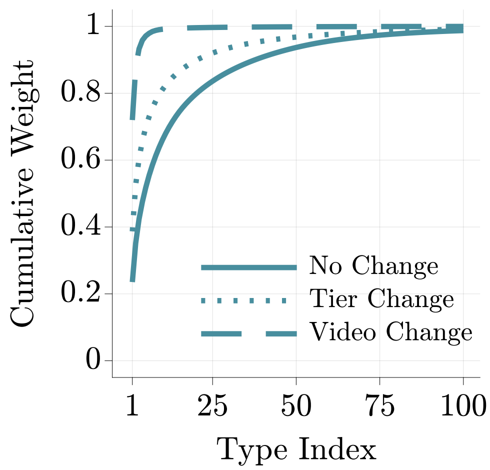
Notes: Summary statistics of the estimated type distribution. Standard errors for each statistic are shown in parentheses and are bootstrapped using 500 samples with replacement of the full set of households used in estimation.

As we discussed in Section 4, we use the information in the data to update the posterior distribution of parameters. The aggregate posterior distributions make clear that the data substantially refine the uniform prior across the R candidate types. One measure of the degree of refinement is the concentration of probability mass within each household's posterior type weights. To obtain this measure, we sort each household's posterior type weights (ω_{hr}) in descending order of magnitude. We then calculate the cumulative weight assigned to all types up to each position in the sorted list. This CDF-style measure shows how many types are needed to describe a given amount of posterior probability mass for a household, or equivalently how much cumulative weight is associated with a given number of types. In Figure 4, we plot the average relationship between number of types and cumulative weight for three mutually exclusive and collectively exhaustive categories of households: those that make no plan changes, those that change their internet tier but not their video plan, and those that change their video plan. For an average household that adds or drops video, the 10 types with the greatest weight have a cumulative value that is close to 100%. Households that change their internet tier have an average of 80% of weight on 10 types, and households with no plan change have an average cumulative posterior weight of 65% on 10 types. This implies that the substantial variation in parameter distributions displayed in Figure A1 is due to heterogeneity across households rather than diffuse posteriors for individual households.

In Appendix B we provide additional illustrations of how variation in the choice environment and data contribute to refining household-level posteriors. We show that posteriors are substantially less concentrated if we use only plan choice (i.e., no usage data) in estimation, or if we use only pre-UBP plan choice and usage data. We also show the impact of reducing a household's choice problem to a single subscription decision in each of the pre-UBP and UBP periods, as well as the impact of narrowing the sample period to the two months on either side of UBP's introduction.

Next, we assess model fit. In Table 3 we compare the empirical and model-predicted market shares and usage levels by plan, where we distinguish six plan categories by speed (Low, Median, High) and internet-only versus bundle (i, b). We compute each value under both empirical pricing models: subscription prices only (pre-UBP) and UBP. Overall, the predicted market shares and usage are comparable to the empirical levels, and key patterns such as movement into higher-allowance tiers during the UBP period are captured by the model. In terms of usage predictions, the model generally matches key empirical usage differences by plan type, including the large difference in levels between internet-only and bundle households, and the increasing average usage across tiers by speed, both before

Figure 4: Household-level Posterior Refinement



Notes: Each household's posterior type weights are sorted in descending order of magnitude, then the weight attached to each position in the sorted vector is averaged across households. Households are divided into three mutually-exclusive, collectively-exhaustive groups based on observed subscription choices. Each curve is interpretable as a CDF showing the average cumulative weight across the 100 highest-weight types for each household.

Table 3: Model Fit

Shares	Pre-UBP Period Model	Pre-UBP Period Data	UBP Period Model	UBP Period Data
i Low Speed	7.8	7.2	7.0	7.3
i Median Speed	13.1	12.0	11.2	11.6
i High Speed	3.0	3.5	4.5	4.6
b Low Speed	18.7	18.6	18.1	17.9
b Median Speed	49.6	51.1	48.6	48.8
b High Speed	7.9	7.6	10.5	9.8
	Pre-UBP Period		UBP Period	
Mean Usage	Model	Data	Model	Data
i Low Speed	144.9	95.0	83.4	102.8
i Median Speed	169.4	169.5	125.5	173.0
i High Speed	204.4	272.5	212.2	334.2
b Low Speed	60.3	41.4	43.9	45.0
b Median Speed	98.0	83.7	75.0	85.4
b High Speed	141.9	149.2	167.7	188.3

Notes: Empirical and model-predicted market shares and average monthly usage levels by plan. Plans are grouped by bundle status (i for internet-only and b for bundle) and internet download speed (Low, Median, High).

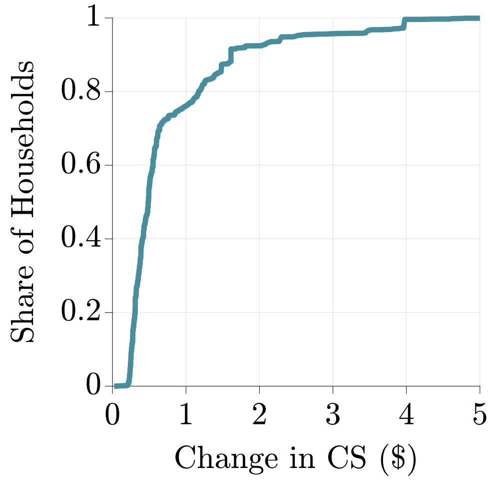
and during the UBP period. As one might expect, the model matches the empirical usage levels most accurately for the highest-share tiers where we have more information (e.g., b Median Speed, which is chosen by nearly half of the households in the data). The model is less accurate for plans with very low take-up, but these contribute relatively little to the overall type distribution (e.g., i High Speed has less than 5% market share).

5.2 Consumer Surplus

We use our estimates to compute willingness-to-pay (WTP) for various aspects of the internet service. The first measure we present is the WTP for greater connection speeds. We compute this WTP as the difference in consumer surplus for households facing the observed menu of internet plans versus a menu with the same prices but internet speeds one megabit per second faster for all tiers.[15] For a given menu of internet plans with price

[15]Across plans, the average observed speed is about 48Mbps.

Figure 5: Willingness to pay for 1 Mbps Speed Increase



Notes: Distribution of the change in consumer surplus (measured in dollars) resulting from a 1 Mbps increase in the download speed associated with all internet tiers.

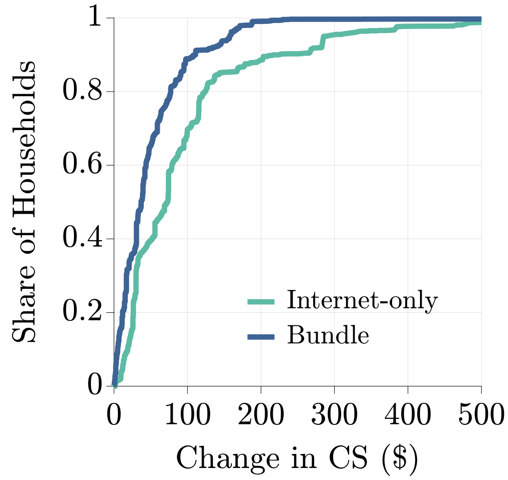
schedules $\boldsymbol{\mathcal{P}}$ and speeds $\mathbf{s}$, each indexed by k, type r's consumer surplus is

$$CS_r(\boldsymbol{\mathcal{P}}, \mathbf{s}) = \frac{1}{\alpha_r} \log \left[\sum_k \exp \left(-\frac{\phi_r}{s_k} + w^*(\theta_r, \mathcal{P}_k) - \alpha_r \left(f_k + \mathcal{O}^*(\theta_r, \mathcal{P}_k) \right) \right) \right].$$

For a household of type r, the WTP for internet speed is the difference in CS_r values for the observed and augmented speed values. We use the estimated distribution of types, ω_r, to compute the distribution of CS_r values in the population, and we repeat this exercise for both empirical price schedules (pre-UBP and UBP). In Figure 5 we present the willingness-to-pay for the speed increase for one complete billing cycle under pre-UBP pricing. The distribution of ϕ, together with the functional form of s_k in utility, generate marginal valuations that are large but decline rapidly for tiers with greater speeds. The average consumer surplus change is \$0.81 per month with pre-UBP pricing and \$0.73 per month with UBP pricing. For either price schedule, within a given plan usage does not change with s_k, so the difference in surplus changes must follow from how the consumers view the utility differences across menu options.

Next, we describe the dollar value households place on access to OTT. For type r, this value is equal to $(\delta_r v_{r,2}/\alpha_r)$, and we compute its distribution using the posterior weights across candidate types. In Figure 6 we present the cumulative distribution function of this dollar value, both for types that prefer internet-only plans and for types that prefer

Figure 6: Willingness to pay for OTT Video

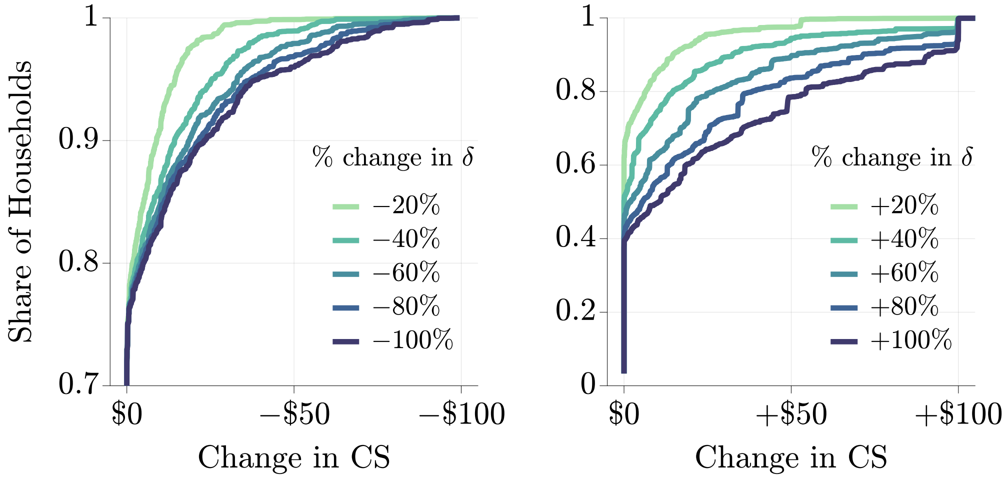


Notes: Marginal distribution of $\delta v_2/\alpha$. The Internet-only distribution weights type-specific values using both the estimated type weights and the probability that the type chooses an internet-only plan, while the Bundle distribution weights type values using the probability that the type chooses an internet and TV bundle plan.

bundle plans. Both distributions are right-skewed, with substantial variation. Types that prefer internet-only plans have a 25th percentile valuation of $29 per month, median of $73, and 75th percentile of $115. Types that prefer a bundled internet and TV plan intuitively have lower OTT valuations, with a 25th percentile of $15, median of $37, and 75th percentile of $66.

Finally, we consider the impact of changing δ on consumer surplus. Following a similar approach to measuring WTP for connection speed increases, we calculate the difference in consumer welfare between the status quo and proportional changes to δ's value. In our model, when households prefer the bundle, a change in δ has no effect on welfare, holding subscription choices fixed. A change in δ affects utility conditional on an internet-only subscription, and therefore can affect the relative utility of internet-only versus the bundle. In Figure 7, we present the distribution of consumer welfare changes for decreases (left panel) and increases (right panel) in δ, holding prices fixed at pre-UBP levels. When δ is reduced, roughly 70% of households are not impacted meaningfully because they would have selected the bundle. Others minimize their losses by choosing the bundle rather than internet-only. If δ is set to zero (i.e., OTT access is completely eliminated), about 15% of households lose $20 or more in welfare, and 5% lose more than $40.

Figure 7: The Effect of a Change in δ on Consumer Surplus



Notes: We show the dollar change in consumer surplus from changes in δ. Each consumer's estimated value, δ_0, is scaled by a proportional factor.

If δ improves, on the other hand, a larger proportion of households can benefit because consumers become more likely to subscribe to an internet-only plan. If each household's δ improves by 40% (capped at 1), the mean consumer welfare increase is \$17, with 15% of households gaining more than \$20 and 10% gaining more than \$40. In Appendix Figure A5 we present analogous results for when the MSO uses UBP. The results are qualitatively very similar with only slight changes in the magntitudes of the welfare implications. For example, the benefits to consumers from an increase in δ are smaller because some of the gains are captured by the MSO through UBP overage charges.

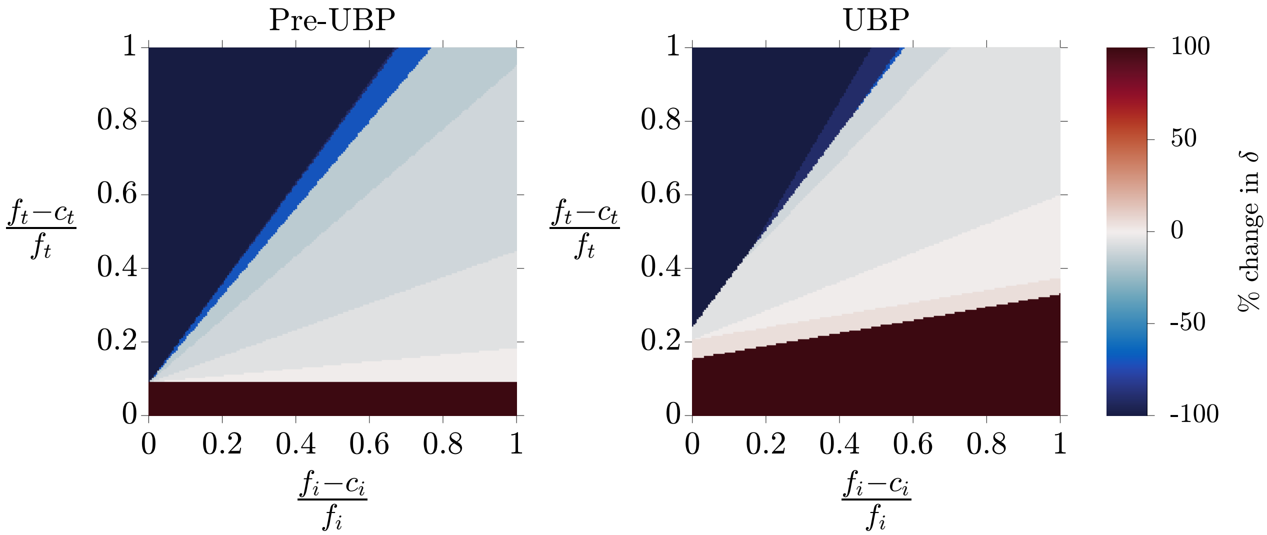
5.3 The MSO's Steering Incentives

As we discussed in Section 2.3, the MSO might have an incentive to steer consumers by altering δ or by setting internet usage prices. We explore both of these strategies using our estimates.

First, we consider the MSO's incentive to directly impact δ, which the MSO might accomplish through a variety of channels. For instance, in its interactions with customers, the MSO could increase δ by providing subsidized or free streaming hardware, or decrease δ by customizing the hardware to limit access to particular OTT content providers. Upstream from consumers, the MSO could alter its network investment, which could either facilitate or impede access to high-quality streaming video. In the absence of net neutrality regulation, an MSO may simply throttle certain sources of traffic like OTT.

We explore the MSO's incentives to alter δ under different assumptions about the

Figure 8: Profit-maximizing change in δ



Notes: The level of the heat map corresponds to a percent change in all households' δ levels, e.g., at 1, δ is increased by 100% (capped at $\delta = 1$), at 0, δ remains at the baseline level, and at -1, all δ values are reduced by 100% down to zero for different levels of costs (c_i and c_t), holding subscription prices (f_i and f_t) at the empirical levels.

relative profitability of internet and video. Specifically, we assume that the MSO's subscription prices for internet and TV, f_i and f_t, are held constant at their empirical pre-UBP levels, while we consider different values for the firm's internet and TV costs, c_i and c_t. We prefer this approach to inferring marginal costs through assumptions on optimal pricing because MSO pricing decisions are complex due to technological, contracting, and regulatory considerations. This approach also allows us to consider a range of potential cost values to account for differences across MSOs (e.g., due to differences in TV licensing fees), and comment on which values are representative of current industry conditions while providing guidance if future conditions differ.

We assume no cost of changing δ, so our results are best interpreted as directional effects on the circumstances when an MSO would affect δ positively or negatively at the margin. We set marginal costs at the levels that imply different degrees of profitability for the two services. Given the subscription prices and costs, we calculate the (single) proportional change to all households' δ values that maximizes the MSO's profit. We present the results in Figure 8 for the observed pre-UBP pricing (left panel) and UBP pricing (right panel), respectively. Along the axes we display different cost conditions. It is convenient to hold subscription fees fixed at their empirical levels and express the cost changes as a change in the proportion of the fees, $\frac{f_i - c_i}{f_i} = 1 - \frac{c_i}{f_i}$ and $\frac{f_t - c_t}{f_t} = 1 - \frac{c_t}{f_t}$. We assume $c_b = c_i + c_t$. Thus, movement outward along either axis implies lower costs or greater relative profitability. Using numbers from Crawford et al. (2018), publicly-reported values for affiliate and re-transmission fees, and the BLS price index for TV services, we calculate that $\frac{f_t - c_t}{f_t}$ has decreased steadily and is no greater than 40% during our sample.[16]

Under pre-UBP pricing, for given internet costs (c_i) the firm benefits from greater δ when TV costs (c_t) are high, but it prefers a lower δ when c_t is smaller. Lower internet costs decrease the incentive to degrade δ. When the MSO charges usage-based prices, there are fewer cost circumstances when the firm wants to reduce δ, and in more situations the firm would benefit from larger δ. The difference in δ incentives between panels is intuitive: with UBP, the MSO can capture rents associated with δ for households that prefer OTT alternatives to the bundle.

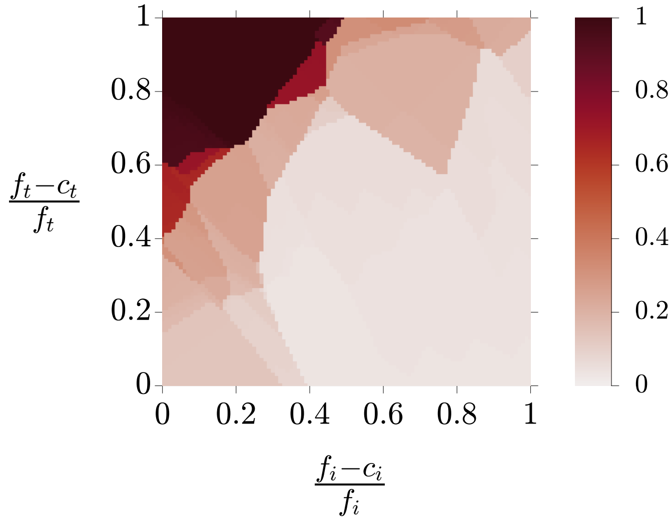
Next, we investigate the MSO's steering incentives using pricing tools and how these incentives vary with the relative profitability of internet and TV. We explore these incentives by studying the impact of differential pricing of OTT and non-OTT internet usage. As we describe in Section 2.3, when studying the pricing incentives of OTT content, one needs to avoid confounding steering and metering incentives. To study the pricing incentives, therefore, we take as given the pre-UBP menu of subscription prices, and we let the MSO choose a new price schedule $\mathcal{P}^{\tau}$, comprised of two prices: τ, which applies uniformly to every unit of internet usage, and τ_δ, which applies to OTT usage only. This fee structure allows the ISP to act on two pricing incentives. First, it can increase the price of internet overall by increasing τ. Second, it can exercise a discriminatory incentive by changing τ_δ to raise or lower the effective price of OTT consumption ($\tau + \tau_\delta$).

In Figure 9, we show how the profit-maximizing OTT usage price, $\tau + \tau_\delta$, varies with the relative profitability of internet and TV. The price is positive for all cost levels, and it is larger as the TV costs decrease. This might be read as suggesting that an MSO wants to disadvantage OTT content and steer consumers to its TV service. However, that conclusion is premature since it is important to decompose the incentives that contribute to $\tau + \tau_\delta$.

In Figure 10, we show both the profit-maximizing τ (left panel) and τ_δ (right panel) levels for each combination of costs. The optimal τ decreases as costs decrease because lost subscriptions are not worth the additional revenue from usage. If subscriptions are less profitable, i.e., near the origin, the MSO prefers a greater value of τ. The MSO's

[16]Price and fee estimates from 2000-2010 are reported in Tables A.I and A.II of Crawford et al. (2018). Publicly-reported values for affiliate and re-transmission fees (the largest components of c_t) are drawn from news articles that reference SNL Kagan data. The BLS price index is "cable, satellite, and live streaming television service" (Series ID CUUR0000SERA02).

Figure 9: Profit-maximizing OTT usage fee

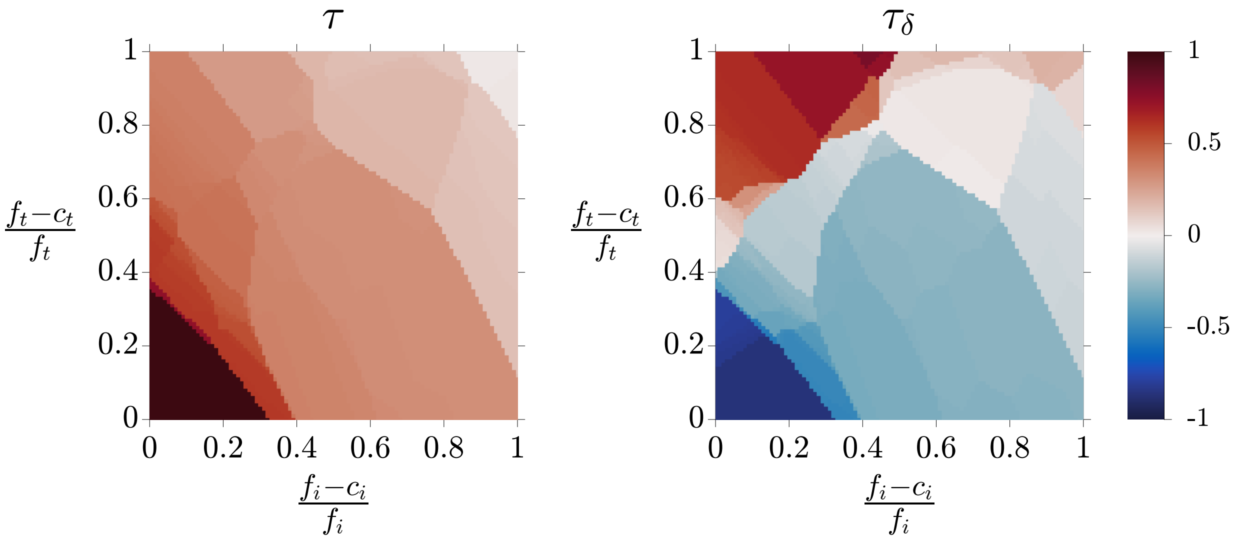


Notes: Profit-maximizing $(\tau + \tau_\delta)$ for different levels of costs (c_i and c_t), holding subscription prices (f_i and f_t) at the empirical levels.

isolated steering incentive is apparent in the optimal values of τ_δ. For a fixed value of internet costs, as TV profitability increases the MSO is more likely to set $\tau_\delta > 0$, steering consumers toward the bundle. Across the full range of TV and internet costs that we consider, charging a positive fee for online video usage (net of the MSO's general incentive to meter internet usage) is only profitable for low TV costs. The preferred pricing policy for relatively high internet and low TV profitability, which is likely more relevant empirically, is to discount OTT relative to non-video internet usage. Reducing τ_δ helps induce bundle subscribers to cut the cord, and the MSO trades a (low) TV margin for metered OTT use at the price $\tau + \tau_\delta$. In Appendix B, we present the consumer welfare implications of the profit-maximizing usage fees.

In Table 4, we present details on the implications of optimal linear usage fees for sources of MSO revenue and changes in household subscriptions and welfare relative to the bundle-pricing baseline. Each column of Table 4 presents these values for different combinations of the relative profitability of TV and internet, $\left(\frac{f_i - c_i}{f_i}, \frac{f_t - c_t}{f_t}\right)$. For example, when $\frac{f_i - c_i}{f_i} = 0.5$ and $\frac{f_t - c_t}{f_t} = 0.3$, the MSO sets $\tau^* = 0.325$ and offers a discount of \$0.05 to OTT traffic. Relative to the bundle-pricing baseline, total revenue increases marginally by about \$1 per household, but is now comprised of both fixed fee subscription revenue and usage fees. Average monthly revenue from usage fees equals \$21.76 and \$14.66 for internet-only and bundle subscribers, respectively. This implies a decrease in consumer welfare of \$18.96. We find little substitution into the bundle for any combination of costs,

Figure 10: Profit-maximizing usage fees



Notes: Optimal τ and τ_δ for different levels of costs (c_i and c_t), holding subscription prices (f_i and f_t) at the empirical levels.

reinforcing our finding that UBP has limited impact as a steering instrument for the MSO. Most of the substitution resulting from optimal linear usage fees is to the outside good.

Table 4: Summary of Pricing Results

$\left(\frac{f_i-c_i}{f_i}, \frac{f_t-c_t}{f_t}\right) =$	(0.5,0.1)	(0.5,0.3)	(0.7,0.1)	(0.7,0.3)	(0.9,0.1)	(0.9,0.3)
τ^*	0.355	0.325	0.325	0.320	0.320	0.315
τ^*_δ	-0.320	-0.275	-0.285	-0.280	-0.285	-0.275
$\tau^* + \tau^*_\delta$	0.035	0.050	0.040	0.040	0.035	0.040
Revenue	150.190	150.607	150.631	150.686	150.668	150.726
i Usage Fees	21.287	21.761	21.095	20.990	20.593	20.883
b Usage Fees	15.253	14.664	14.664	14.559	14.559	14.452
Consumer Harm	-20.032	-19.024	-18.958	-18.767	-18.732	-18.574
s_i	0.147	0.150	0.152	0.153	0.154	0.154
s_b	0.717	0.719	0.719	0.719	0.719	0.720
s_0	0.136	0.131	0.129	0.128	0.127	0.127

Notes: Firm and consumer choices under linear usage fee pricing for a range of assumed internet and TV marginal costs. τ^* and τ^*_δ are the optimal linear usage fees holding subscription fees fixed at the empirical levels. Revenue is the sum of subscription fees and usage fees. I and B usage fees are the average fees collected from internet-only and bundle plan subscribers. Consumer harm is the difference between consumer surplus with linear usage fees and consumer surplus with pre-UBP pricing as a baseline. s_i, s_b, s_0 are the choice shares of internet-only plans, bundle plans, and the outside option, respectively.

6 Conclusions

We study the pricing and quality-provision incentives of MSOs, which serve as gatekeepers in providing internet access. OTT increases the demand for an MSO's internet subscription services, but this may come at the expense of reduced subscriptions to an MSO's TV service. In confronting these challenges, an MSO may use prices or direct intervention in OTT quality to benefit from an improved internet offering or steer consumers away from OTT and towards its own TV service.

We specify and estimate a model that captures the central incentives behind MSOs' policies. We show that indeed an MSO might have an incentive to steer consumers away from OTT, but the strength and even the sign of the steering incentive depends on an MSO's relative costs of internet and TV service. In addition, we show that, when an MSO is able to use usage-based pricing for internet content including OTT, it is more likely to benefit from increasing OTT quality.

Understanding the incentive to steer is relevant for antitrust policy in the telecommunications industry. In particular, the evaluation of mergers, whether between distribution firms or between content and distribution firms, presents a number of challenges. First, market boundaries may be difficult for regulators and antitrust authorities to identify because little evidence exists on consumers' willingness to substitute across conventional TV, streaming video, and other non-video internet applications. Our results show that consumers derive substantial value from OTT video. Thus, antitrust analysis might need to consider broad market definitions that encompass many forms of digital entertainment, as well as the central role of MSOs in shaping how content is distributed and surplus is allocated. Second, antitrust authorities need to assess how existing or new vertical relationships may affect an MSO's incentives to introduce restrictive cross-licensing agreements or use price instruments to favor its own content over competitors'. The impact of these strategies depends on consumers' responsiveness to steering. An MSO that is vertically integrated with a content-producing firm may foreclose some content from availability to consumers via a competing MSO.[17] Our estimates show that even blunt mechanisms like usage-based pricing can have important allocative consequences among consumers and various firms.

More broadly, our results are also relevant for the net neutrality debate, in which empirical evidence is rare. Net neutrality's 2017 repeal provides MSOs more latitude to discriminate across types of internet traffic. While we do not observe source-specific discrimination in our data, our results are informative about MSOs' incentives to discriminate when they have the opportunity. For example, MSOs may respond to increased popularity of individual applications by introducing application-specific prices or barriers to extract some of the surplus from OTT innovations.

There are several issues our model and empirical results do not address, and we leave for future research. While our model provides a useful framework for studying the steering incentives of MSOs, a richer specification is required to quantify substitution patterns between specific applications and content providers. Similarly, the model makes simplifying assumptions on the interaction between firms, for example by holding fixed OTT supply. Given the differences in OTT content across applications and the potential pricing power of third-party firms, modeling and evaluating the relationships between these firms and MSOs is a fruitful area for future research.

[17] A price-based steering strategy with similar effects is "zero rating," which favors certain content by not counting its usage against a monthly allowance. Zero rating has been used by telecommunications providers including T-Mobile and Comcast (https://www.vice.com/en/article/nz7nyx/comcast-hit-with-fcc-zero-rating-complaint-over-stream-tv).

References

Ackerberg, Daniel A (2009). "A New Use of Importance Sampling to Reduce Computational Burden in Simulation Estimation." *Quantitative Marketing and Economics*, 7(4): 343–376.

Armstrong, Mark (2006). "Competition in Two-Sided Markets." *RAND Journal of Economics*, 37(3): 668–691.

Armstrong, Mark (2013). "A More General Theory of Commodity Bundling." *Journal of Economic Theory*, 148(2): 448–472.

Barwick, Panle Jia, Parag A. Pathak and Maisy Wong (2017). "Conflicts of Interest and Steering in Residential Brokerage." *American Economic Journal: Applied Economics*, 9(3): 191–222.

Bourreau, Marc, Frago Kourandi and Tommaso Valletti (2015). "Net Neutrality with Competing Internet Platforms." *Journal of Industrial Economics*, 63(1): 30–73.

Choi, Jay Pil, Doh-Shin Jeon and Byung-Cheol Kim (2015). "Network Neutrality, Business Models, and Internet Interconnection." *American Economic Journal: Microeconomics*, 7(3): 104–141.

Choi, Jay Pil and Byung-Cheol Kim (2010). "Net Neutrality and Investment Incentives." *RAND Journal of Economics*, 41(3): 446–471.

Chu, Chenghuan Sean, Phillip Leslie and Alan Sorensen (2011). "Bundle-Size Pricing as an Approximation to Mixed Bundling." *American Economic Review*, 101(1): 263–303.

Cisco (2018). "Cisco Visual Networking Index (VNI) Complete Forecast Update, 2017–2022."

comScore (2018). "State of OTT, 2018."

Crawford, Gregory and Matthew Shum (2007). "Monopoly Quality Degradation and Regulation in Cable Television." *Journal of Law and Economics*, 50(1): 181–219.

Crawford, Gregory and Ali Yurukoglu (2012). "The Welfare Effects of Bundling in Multichannel Television Markets." *American Economic Review*, 102(2): 643–685.

Crawford, Gregory S., Robin S. Lee, Michael D. Whinston and Ali Yurukoglu (2018). "The Welfare Effects of Vertical Integration in Multichannel Television Markets." *Econometrica*, 86(3): 891–954.

Crawford, Gregory S., Oleksandr Shcherbakov and Matthew Shum (2019). "Quality Overprovision in Cable Television Markets." *American Economic Review*, 109(3): 956–95.

Dutz, Mark, Jonathan Orszag and Robert Willig (2009). "The Substantial Consumer Benefits of Broadband Connectivity for US Households." *Internet Intervention Alliance Working Paper*.

Economides, Nicholas and Benjamin Hermalin (2012). "The Economics of Network Neutrality." *The RAND Journal of Economics*, 43(4): 602–629.

Economides, Nicholas and Benjamin Hermalin (2015). "The Strategic Use of Download Limits by a Monopoly Platform." *The RAND Journal of Economics*, 46(2): 297–327.

Economides, Nicholas and Joacim Tag (2012). "Network Neutrality on the Internet: A Two-Sided Market Analysis." *Information Economics and Policy*, 24(2): 91–104.

Economides, Nicholas and Joacim Tag (2016). "Internet Regulation, Two-Sided Pricing, and Sponsored Data." *Working Paper*.

Edell, Richard and Pravin Varaiya (2002). *Providing Internet Access: What We Learn from INDEX*, volume Broadband: Should We Regulate High-Speed Internet Access? Brookings Institution.

Fox, Jeremy T, Kyoo il Kim and Chenyu Yang (2016). "A Simple Nonparametric Approach to Estimating the Distribution of Random Coefficients in Structural Models." *Journal of Econometrics*, 195(2): 236–254.

Fox, Jeremy T, Stephen P Ryan and Patrick Bajari (2011). "A Simple Estimator for the Distribution of Random Coefficients." *Quantitative Economics*, 2(3): 381–418.

Gans, Joshua (2015). "Weak Versus Strong Net Neutrality." *Journal of Regulatory Economics*, 47(2): 183–200.

Gentzkow, Matthew (2007). "Valuing New Goods in a Model with Complementarity: Online Newspapers." *American Economic Review*, 97(3): 713–744.

Goetz, Daniel (2019). "Dynamic Bargaining and Size Effects in the Broadband Industry." *Working Paper*.

Goolsbee, Austan and Peter Klenow (2006). "Valuing Products by the Time Spent Using Them: An Application to the Internet." *American Economic Review P&P*, 96(2): 108–113.

Greenstein, Shane and Ryan McDevitt (2011). "The Broadband Bonus: Estimating Broadband Internet's Economic Value." *Telecommunications Policy*, 35(7): 617–632.

Greenstein, Shane, Martin Peitz and Tommaso Valletti (2016). "Net Neutrality: A Fast Lane to Understanding the Tradeoffs." *Journal of Economic Perspectives*, 30(2): 127–150.

Grubb, Michael (2015). "Consumer Inattention and Bill-Shock Regulation." *Review of Economic Studies*, 82(1): 219–257.

Grubb, Michael and Matthew Osborne (2015). "Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock." *American Economic Review*, 105(1): 234–271.

Hagemann, Garrett (2017). "Upstream Quantity Discounts and Double Marginalization in the New York Liquor Market." *Working Paper.*

Hitte, Loran and Prasanna Tambe (2007). "Broadband Adoption and Content Consumption." *Information Economics and Policy*, 74(6): 1637–1673.

Ho, Kate and Robin Lee (2019). "Equilibrium Insurer-Provider Networks: Bargaining and Exclusion in Health Care Markets." *American Economic Review*, 109(2).

Lambrecht, Anja, Katja Seim and Bernd Skiera (2007). "Does Uncertainty Matter? Consumer Behavior Under Three-Part Tariffs." *Marketing Science*, 26(5): 698–710.

Lee, Kwok Hao and Leon Musolff (2021). "Entry Into Two-Sided Markets Shaped By Platform-Guided Search." *Working Paper.*

Lee, Robin and Tim Wu (2009). "Subsidizing Creativity through Network Design: Zero-Pricing and Net Neutrality." *Journal of Economics Perspectives*, 23(3): 61–76.

Liebman, Eli (2017). "Bargaining in Markets with Exclusion: An Analysis of Health Insurance Networks." *Working Paper.*

Malone, Jacob, Aviv Nevo and Jonathan Williams (2016). "A Snapshot of the Current State of Residential Broadband Networks." *NET Institute Working Paper No. 15-06.*

Malone, Jacob, Aviv Nevo and Jonathan W. Williams (2021a). "The Tragedy of the Last Mile: Economic Solutions to Congestion in Broadband Networks." *Working Paper.*

Malone, Jacob, John Turner and Jonathan Williams (2014). "Do Three-Part Tariffs Improve Efficiency in Residential Broadband Networks?" *Telecommunications Policy*, 38(11): 1035–1045.

Malone, Jacob B., Aviv Nevo, Zachary Nolan and Jonathan W. Williams (2021b). "Is OTT Video a Substitute for TV? Policy Insights from Cord-Cutting." *The Review of Economics and Statistics (forthcoming).*

MarketWatch (2018). "Over 5 million U.S. consumers will cut the cord in 2018, survey says."

McManus, Brian (2007). "Nonlinear Pricing in an Oligopoly Market: the Case of Specialty Coffee." *RAND Journal of Economics*, 38: 512–532.

McManus, Brian, Aviv Nevo, Zachary Nolan and Jonathan W. Williams (2023). "The Implications of Usage-based Pricing in Telecommunications." *Work in Progress.*

Miravete, Eugenio (2003). "Choosing the Wrong Calling Plan? Ignorance and Learning." *American Economic Review*, 93(1): 297–310.

Mussa, Michael and Sherwin Rosen (1978). "Monopoly and Product Quality." *Journal of Economic Theory*, 18(2): 301–317.

Nevo, Aviv, John Turner and Jonathan Williams (2016). "Usage-Based Pricing and Demand for Residential Broadband." *Econometrica*, 84(2): 411–443.

Prince, Jeffrey and Shane Greenstein (2017). "Measuring Consumer Preferences for Video Content Provision via Cord-Cutting Behavior." *Journal of Economics & Management Strategy*, 26(2): 293–317.

Raval, Devesh (2022). "Steering in One Click: Platform Self-Preferencing in the Amazon Buy Box." *Working Paper*.

Raval, Devesh and Ted Rosenbaum (2017). "Payer Steering and Patient Welfare: Evidence from Medicaid." *Working Paper*.

Reggiani, Carlo and Tommaso Valletti (2016). "Net Neutrality and Innovation at the Core and at the Edge." *International Journal of Industrial Organization*, 45(1): 16–27.

Rosston, Gregory, Scott Savage and Bradley Wimmer (2013). "Effect of Network Unbundling on Retail Price: Evidence from the Telecommunications Act of 1996." *Journal of Law and Economics*, 56(2): 487–519.

Sidak, Gregory (2006). "A Consumer-Welfare Approach to Network Neutrality Regulation of the Internet." *Journal of Competition Law & Economics*, 2(3): 349–474.

Tudon, Jose (2021). "Prioritization vs. congestion on platforms: Evidence from Amazon's Twitch.tv." *Working Paper*.

Varian, Hal (2002). *The Demand for Bandwidth: Evidence from the INDEX Experiment*, volume Broadband: Should We Regulate High-Speed Internet Access? Brookings Institution.

Wu, Tim (2003). "Network Neutrality, Broadband Discrimination." *Journal of Telecommunications and High Technology Law*, 1(2): 141–178.

Appendix

A The Empirical Price Schedule

In this section, we provide additional detail on the empirical UBP schedule and its implications for consumers' usage choices. The full expression for UBP payments is:

$$\mathcal{O}(q_1, q_{2,i}; \mathcal{P}_k^e) = \begin{cases} 0 & \text{if} \quad q \leq \kappa \\ \overline{p}_k & \text{if} \quad \kappa < q \leq \kappa_k + \overline{q}_k \\ 2\overline{p}_k & \text{if} \quad \kappa + \overline{q}_k < q \leq \kappa_k + 2\overline{q}_k \\ \quad\vdots & \\ n\overline{p}_k & \text{if} \quad \kappa + (n-1)\overline{q}_k < q \leq \kappa_k + n\overline{q}_k. \end{cases}$$

We capture the possibility of an arbitrary number of top-ups in a more general expression for optimal internet usage by bundle subscribers:

$$q_1^*(\mu; \theta, \mathcal{P}_k^e) = \begin{cases} \mu v_1 & \text{if} \quad \mu v_1 > \kappa_k \quad \& \quad v_1 - \frac{1}{\mu}\left(\kappa_k + \overline{q}_k \lfloor \frac{v_1 - \kappa_k}{\overline{q}_k} \rfloor\right) > \alpha \overline{p}_k \\ & \quad \text{or} \quad \mu v_1 \leq \kappa_k \\ \kappa_k + \overline{q}_k \lfloor \frac{v_1 - \kappa_k}{\overline{q}_k} \rfloor & \text{if} \quad \mu v_1 > \kappa_k \quad \& \quad v_1 - \frac{1}{\mu}\left(\kappa_k + \overline{q}_k \lfloor \frac{v_1 - \kappa_k}{\overline{q}_k} \rfloor\right) \leq \alpha \overline{p}_k, \end{cases} \tag{6}$$

where $\lfloor \cdot \rfloor$ is the floor function, i.e. the largest integer less than or equal to the function's argument.[18] Optimal usage results in bunching that generates the mass points in the quantity density $g_1(q; \theta, \mathcal{P}_k^e)$ discussed above.

The structure of the internet-only household's choice follows the same logic as in the bundled household, but with $v_1 + \delta v_2$ replacing all instances of v_1 in equation (6). Overage charges associated with optimal consumption of internet and video content equal

$$\mathcal{O}(q_1^*(\mu; \theta, \mathcal{P}_k^e), q_{2,i}^*(\mu; \theta, \mathcal{P}_k^e); \mathcal{P}_k^e) = \begin{cases} \overline{p}_k \lceil \frac{q_1^*(\mu;\theta,\mathcal{P}_k^e) + q_{2,i}^*(\mu;\theta,\mathcal{P}_k^e) - \kappa_k}{\overline{q}_k} \rceil & \text{if plan } k \text{ is internet-only} \\ \overline{p}_k \lceil \frac{q_1^*(\mu;\theta,\mathcal{P}_k^e) - \kappa_k}{\overline{q}_k} \rceil & \text{otherwise.} \end{cases}$$

where $\lceil \cdot \rceil$ refers to the ceiling function applied to internet usage in excess of the allowance,

[18] A consumer who buys n top-ups will use all of the n^{th} top-up if $\mu v_1 > \kappa_k + n\overline{q}_k$. If this inequality is reversed, the consumer will choose $q_1^* = \mu v_1$ and leave some of the top-up unused. Similar logic applies to a bundle subscriber's total usage.

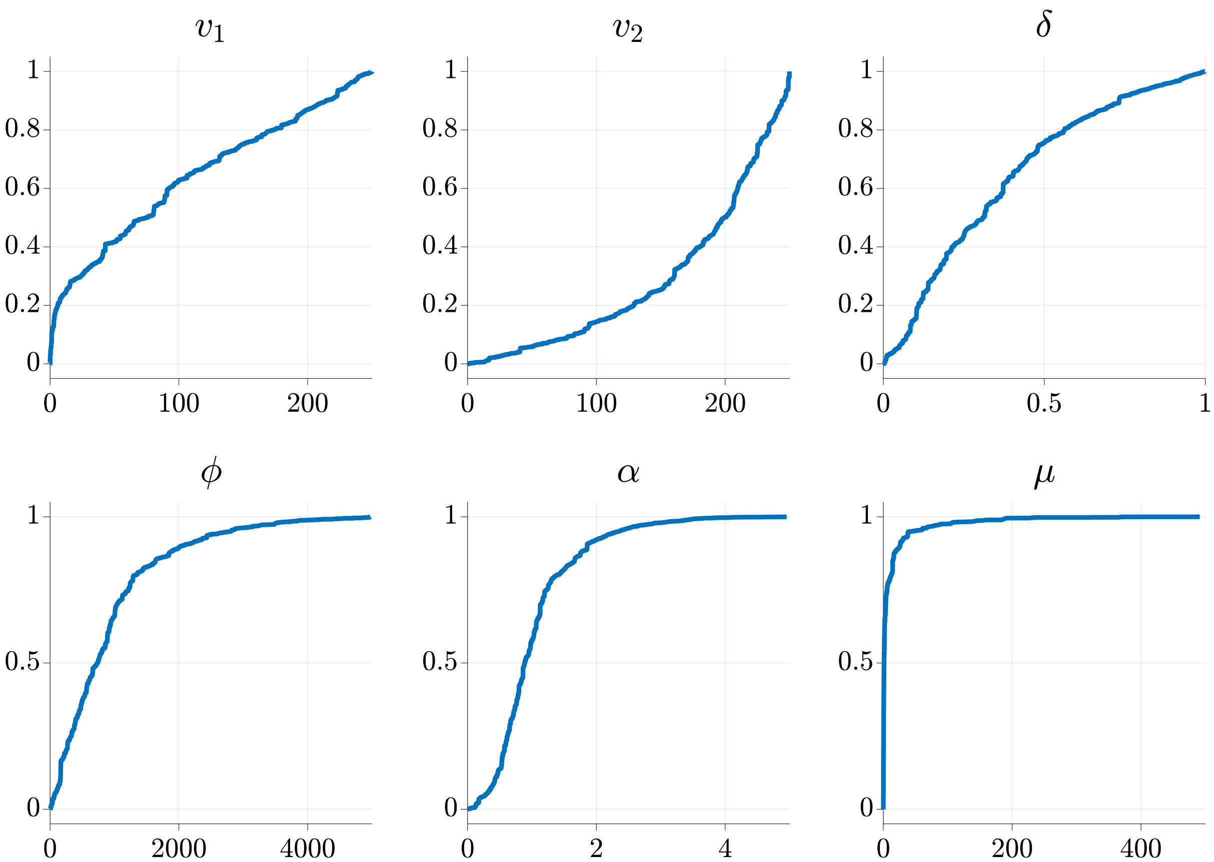
i.e., the smallest non-negative integer greater than or equal to its argument.

B Additional Estimation Results

B.1 Distributions of Parameter Estimates

In Figure A1 we provide graphical displays of each structural parameter's marginal cumulative distribution function.

Figure A1: Marginal Distributions of Structural Parameters



Notes: Estimates of the marginal distributions of the six structural parameters in the empirical model.

B.2 Alternative Likelihood Specifications

In this section, we consider alternative likelihood specifications. These alternatives provide additional intuition on how the different parts of the model and full data identify the preference parameters.

To disentangle the importance of the UBP price change, household usage information, and household plan changes in refining household-level posteriors, we construct the same refinement measure shown in Figure 4 for three alternative household likelihood calculations. Recall that the likelihood function we use ($\mathcal{L}_h(\theta_r)$) includes monthly plan choice and usage decisions both before and after the implementation of UBP. In Figure A2, we show the degree to which the posterior is refined when the likelihood function contains only plan choice information ("No Usage"), and using only the pre-UBP data ("No UBP"). We do this separately for the three classes of consumers highlighted in Figure 4: those with no subscription change during the sample, those who added or dropped TV, and those who changed their internet tier. Our main specification is labeled "Usage; UBP". Intuitively, more information leads to a more concentrated posterior: "No Usage; No UBP" leads to the least-refined posteriors while "Usage; UBP" leads to the greatest degree of refinement. However, the relative importance of the UBP price variation and usage information is household-specific, depending on the household's decisions. For households that made a plan change, plan change information alone adds more information to the "No Usage; No UBP" baseline than usage in the pre-UBP period, while the opposite is true for households that did not make a plan change. Thus, plan changes place meaningful restrictions on the set of types that rationalize household behavior, but even in the absence of plan changes, both the complexity of the subscription menu and the usage choices allow for substantial refinement of the uniform prior.

In a second set of alternative approaches to estimation, we explore the impact of repeated subscription choices across the sample period. We first reduce the number of plan decisions from monthly to a single choice pre-UBP and a single choice post-UBP. The goal of this alternative is to check the robustness of our estimates to unmodeled dimensions of plan choice such as inertia and switching costs which may reduce the true frequency with which a household considers whether to change plans. In this specification, we include all monthly usage decisions, but only the first pre-UBP plan choice and the last plan choice under UBP. We refer to this specification as "Reduced T" below.

In an additional alternative, we reduce the number of time periods used in estimation down to two months on either side of the price change. The goal of this alternative is to check the robustness of our estimates to the assumption that demand parameters can be held constant across time. We refer to this specification as "Reduced Discrete Choice" below.

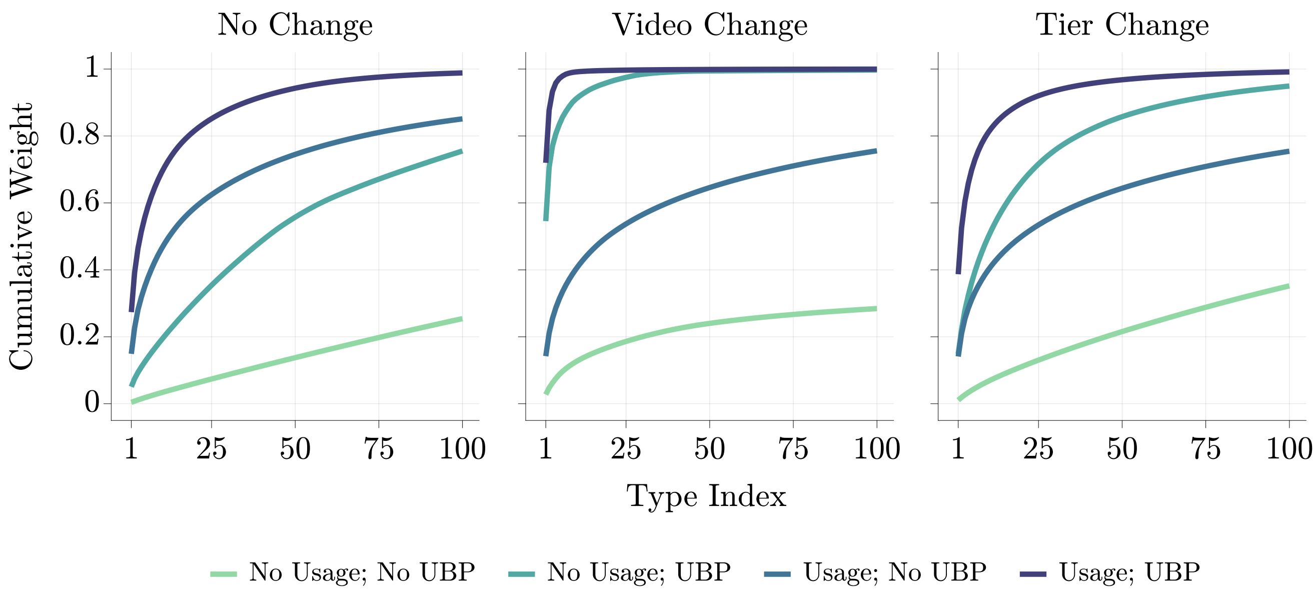
Table 5: Mean Parameter Estimates, Alternative Likelihoods

Likelihood	v_1	v_2	δ	ϕ	λ	α
(1) No Usage; No UBP	114.03	173.58	0.30	1058.25	1.08	24.22
(2) No Usage; UBP	74.17	180.15	0.31	952.48	1.06	35.71
(3) Usage; No UBP	86.20	177.85	0.34	1000.34	1.04	12.01
(4) Usage; UBP	88.27	178.83	0.35	951.13	1.04	10.57
(5) Reduced Discrete Choice	81.30	169.12	0.37	862.93	0.93	12.23
(6) Reduced T	83.46	177.05	0.35	937.89	1.02	13.15

Notes: Mean levels of estimated marginal parameter distributions under alternative likelihood functions.

Table 5 shows the mean of each structural parameter when different likelihood functions are used to generate the posterior. Specification (4) generates our main results, specifications (1)-(3) are described in Section 5, and the last two specifications are described above. Qualitatively, the parameter estimates in the two specifications proposed above are not dissimilar from the specification used to generate our results. The largest changes to the estimates arise when usage information is ignored entirely.

Figure A2: Posterior Refinement, Alternative Likelihood Functions



Notes: Households are divided into three mutually-exclusive, collectively-exhaustive groups based on observed subscription choices. Each curve is interpretable as a CDF showing the average cumulative weight across the 100 highest-weight types for each household.

Figure A3: Posterior Refinement, Reduced T Likelihood

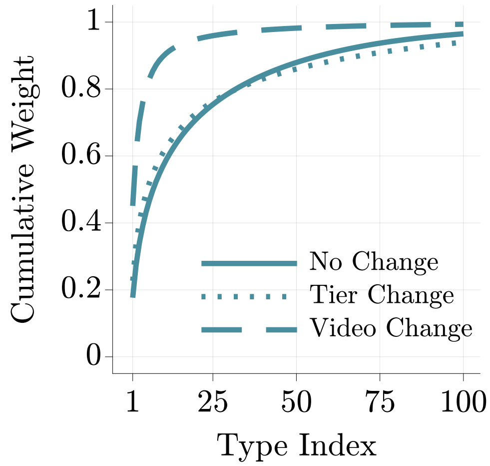
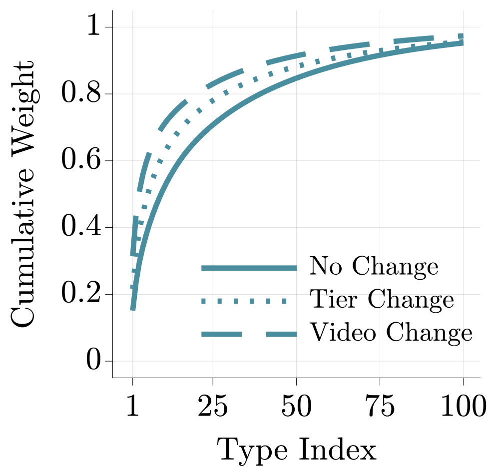


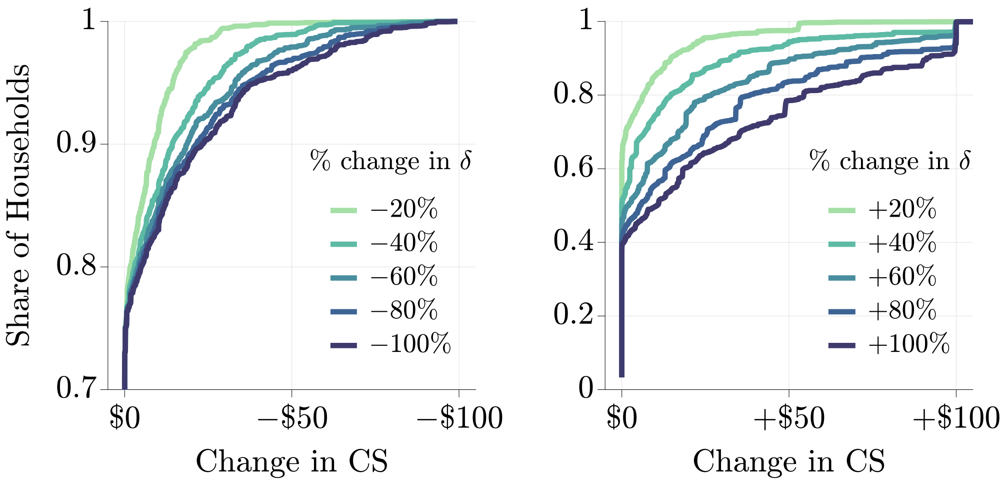
Figure A4: Posterior Refinement, Reduced Discrete Choice Likelihood



B.3 Supplemental Results on Consumer Surplus

Figure A5 presents the analogous results to Figure 8 in Section 5, but with a baseline of UBP as implemented by the MSO rather than the pre-UBP pricing.

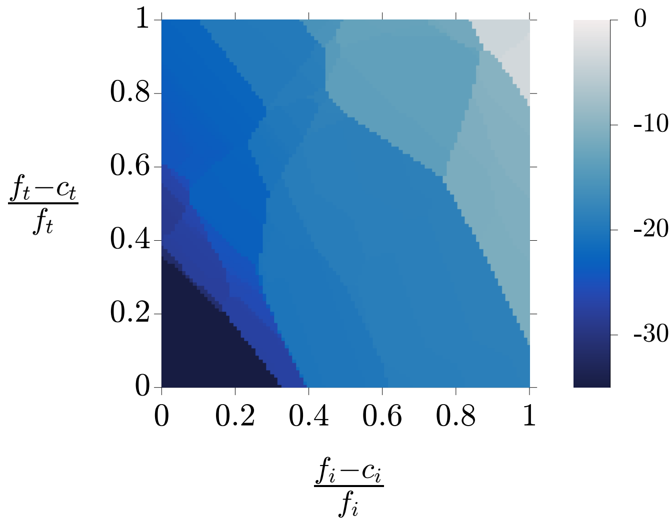
Figure A5: Consumer Surplus Implications of δ Changes (UBP)



Notes: This figure depicts the change in consumer surplus (measured in dollars) resulting from changes in δ. To perform the calculation, each consumer's estimated delta level (δ_0) is scaled by a proportional factor, ranging from 2 down to 0.

Figure A6 depicts the consumer welfare implications from the profit-maximizing linear usage fees presented in Figures 9 and 10.

Figure A6: Consumer surplus at optimal usage fees



Notes: Consumer surplus at optimal τ and τ_δ for different levels of costs (c_i and c_t), holding subscription prices (f_i and f_t) at the empirical levels.