Measuring Consumer Switching Costs in the Television Industry *

Oleksandr Shcherbakov†

Abstract

In this article I develop and estimate a model of dynamic consumer behavior with switching costs in the market for paid television services. I estimate the parameters of the structural model using data on cable and satellite systems across local U.S. television markets over the period 1992-2006. The results suggest switching costs range from $159 to $242 for cable and from $212 to $276 for satellite providers in 1997 dollars. Using a simple dynamic model of cable providers I demonstrate that switching costs of these magnitudes can significantly affect the firms' optimal strategies.

1 Introduction

Switching costs are likely to exist in a variety of important industries, including computer software and hardware development, banking, and telecommunications. However, how consumer switching costs affect competition is not clear from a theoretical standpoint. Do markets become more competitive due to producers' incentives to invest in their customer base? When does the "harvesting" motive (from already locked-in consumers) begin to dominate? Should such industries be regulated and, if so, then when and how? A careful analysis of market structure and industry conduct and design of an efficient policy environment, requires not only a qualitative assessment but also a precise quantitative measure of switching costs and their effects on market outcomes.

In this article I quantify switching costs in the paid television industry and explore potential implications for competition. My primary contribution is the development of an

*I am deeply indebted to my supervisors Gregory Crawford and Gautam Gowrisankaran for their help at all stages of my work. I greatly acknowledge valuable advice and suggestions made by Daniel Ackerberg, Keisuke Hirano, and Aviv Nevo. Useful comments and suggestions made by Kathleen Nosal, Yuya Takahashi, Tim Lee, Nicolas Schutz, Philipp Schmidt-Dengler, and anonymous referees are greatly appreciated. All errors are my own.

†University of Mannheim; ashcherb@mail.uni-mannheim.de

empirical framework that can be used to identify and estimate consumer switching costs using market-level data. I then use estimates of the structural parameters and a supply side model to illustrate differences in the optimal policy of cable television providers under alternative market structures.

In many markets consumers make repeated purchases from the same provider because switching to an alternative supplier involves additional monetary and utility costs. When these costs are substantial and product characteristics change rapidly over time, rational consumers should recognize the effects of their current decisions on their future utility flows. Therefore, allowing for forward-looking behavior of consumers is important for an internally consistent empirical model. For example, in a model where consumer expectations about future product attributes are ignored, consumer decisions can be rationalized only by current period variables. This may result in biased estimates of the structural parameters and incorrect policy implications.

In this article I develop a method to identify switching costs from aggregate data. When individual decisions are not observed, a researcher has to separately identify both the distribution of persistent consumer preferences and the switching costs from aggregate statistics. As in static decision models, I use the variation in observable product characteristics across markets to identify parameters of the distribution of consumer heterogeneity. One way to help identify these parameters is if some consumer decisions are not directly affected by the switching costs. For example, if a consumer can switch between alternative products from one supplier at no cost, product-specific market shares provide important information about the distribution of preferences. By contrast, I identify the switching cost parameters using exogenous shifters of the previous period's consumer decisions because in a model without switching costs (or other sources of state dependence), previous period decisions must be irrelevant for current period choices.

Developments in the U.S. paid television industry between 1992 and 2006 make it a good setting to study consumer switching costs incurred when changing service providers. Although cable television in the United States has a fairly long history, entry of new satellite technology in the 1990s challenged the previous monopoly position of the cable providers. The new market structure provides an opportunity to explore the interplay between consumer switching costs and competition in an important industry.[1]

To empirically evaluate the effect of switching costs on the optimal policy by the cable providers, it is crucial to consistently estimate consumer switching costs and other parameters of the consumer utility function, which is the main focus of the empirical application in this article. The empirical model of consumer decisions I use to quantify consumer switching costs is tailored to accurately represent the institutional details of the paid television industry in the United States, including differences in technologies used to deliver TV signals and vertical differentiation of products offered by the same firm.[2]

To motivate estimation of the structural model, I provide reduced-form evidence of state dependence in consumer decisions. In particular, I find that the current period values of exogenous variables affect not only contemporaneous consumer decisions but also future decisions, which is consistent with substantial consumer switching costs in the industry.

The results suggest that consumer switching costs amount to approximately \$190 for cable and \$240 for satellite (in 1997 dollars). These correspond to slightly more than one-half of the annual service cost for each of the providers. Comparing these estimates to the average cable service installation fees in 1992-2002 implies that the fees account for about 20 percent of the estimated monetary value of switching costs, with the remaining 80 percent of the costs explained by the unobserved hassle costs of switching. Importantly, both static and myopic models of consumer behavior overestimate short-run price elasticity because they ignore the effect of price increases in the current period on the future flow utility values.

To illustrate the importance of accounting for consumer switching costs I develop a dynamic model of a cable service provider and recover the cost structure of cable firms. Then I use the model together with estimates of the demand side parameters to evaluate the effect of consumer switching costs on the optimal policy of cable providers under both

[1] In 2007, cable operators had revenues of about \$79.1 billion nationwide.

[2] The methodology can be applied to other industries as well. For example, Nosal (2012) extends the method to estimate consumer switching costs in the health insurance industry.

duopoly and monopoly scenarios. Although this is not meant to be suggestive of actual policies, it does suggest that the magnitudes I estimate would be material to firms' optimal strategies. For example, holding prices and quality levels by satellite operators fixed and assuming that the cost structure of the cable providers is constant across alternative scenarios, without consumer switching costs and, hence, without incentives to invest in the customer base, cable providers would offer 40 percent lower mean utility (38 percent higher prices) at the industry steady state. The investment incentive effect itself is comparable to the effect of satellite entry on the equilibrium policy by cable firms. Without a satellite competitor but with consumer switching costs in place, prices for cable television would increase by about 48 percent, delivering about 51 percent less consumer utility. These findings show that consumer switching costs may have a nontrivial effect on the market equilibrium. Furthermore, estimating these costs within an internally consistent dynamic model of consumer behavior are crucial for antitrust regulation and a wide range of other policy-related questions.

This article contributes to several strands of the empirical industrial organization literature. First, it provides an identification strategy and a method to estimate consumer switching costs using only aggregate market-level data. Second, it supplements the literature on dynamic demand estimation by accounting for the forward-looking consumer decisions induced by switching costs (an overview of the related theoretical literature can be found in Farrell and Klemperer, 2007). Third, the article empirically evaluates the effects of the switching costs and entry by satellite service providers on the optimal policy of cable incumbents in the paid-television industry.

The literature related to switching costs includes Dube et al. (2009, 2010), who estimate switching costs within a myopic consumers framework using individual-level data. Sudhir and Yang (2014) exploit peculiarities of the individual-level observations (choice-consumption mismatch) to disentangle structural state dependence and unobserved persistent heterogeneity in preferences. Handel (2013) identifies inertia in consumer decisions by using redesign and forced active re-enrollment into health insurance plans. In a myopic consumer model, Yang (2010) shows that switching costs can be identified with market-level data when churn rates are observed. In a dynamic environment, switching costs are studied in Kim (2006), Schiraldi (2011), and Nosal (2012). This article differs in several dimensions, including its identification argument and less restrictive assumptions on the consumer belief about the evolution of state variables (for a detailed discussion of this issue see Aguirregabiria and Nevo, 2013).

Solutions to the consumer dynamic programming problem and the estimation algorithm in this article build on recent research by Hendel and Nevo (2006), Melnikov (2013), and Gowrisankaran and Rysman (2012). Several recent studies address various economic questions using data from the paid-television industry in the United States, including Crawford (2000, 2005), Goolsbee and Petrin (2004), Crawford and Shum (2007), and Crawford and Yurukoglu (2012).

The rest of the article is organized as follows. Section 2 outlines the institutional environment in the U.S. television industry, and the model of consumer behavior is presented in Section 3. Data description, instrumental variables, identification strategy, and reduced-form evidence of state dependence in consumer decisions are provided in Section 4. The estimation algorithm is discussed in Section 5. Section 6 discusses estimation and results. Counterfactual simulations are provided in Section 7 and Section 8 concludes.

2 Television industry: institutional details

Cable television originated in the late 1940s as a mean of delivering broadcast signals to areas with poor over-the-air reception. It diffused widely in the 1970s when television networks began using satellite technology to deliver their content to cable systems. Until the early 1990s, when a Direct Broadcast Satellite (DBS) service was launched, local cable systems were natural monopolies. Since then the subscriber base of DBS providers has experienced rapid growth. Competition in the television industry between DBS and cable operators is somewhat unusual because cable systems make pricing and quality decisions locally, whereas satellite policies are set at the national level. In 1996, the

Telecommunications Act removed price controls on high-quality products leaving only basic service subject to (possibly very weak) regulation. So-called "must-carry" regulations sometimes require cable systems to provide certain television stations in their area. Beyond these restrictions cable providers have complete control over the content and price.[3]

In the paid-television industry, products are vertically differentiated bundles of programming networks with higher quality more expensive bundles uniformly including all the channels from the low quality bundles. This feature of the industry is used to construct a scalar variable measuring quality of the programming content.

Consumers of the paid-television products face substantial switching costs when changing service providers. The most obvious components of switching costs are upfront installation fees and, sometimes, equipment purchases.[4] For instance, in 1997 the costs for the basic equipment, installation, and one month of programming range from $185 for *PrimeStar* satellite service, for which the consumer rents equipment, up to $379 for *DirecTV* service (including $199 equipment, $150 professional installation, and monthly charges of $29.99 for the basic programming package). Average professional installation of cable service was about $40. In addition to the equipment costs there exist hassle costs associated with the choice of and connection to a provider. For example, to arrange equipment installation DBS and cable providers offer a time window for installation appointments during business hours. A customer may also need to get a landlord's permission to install a satellite dish or to wire a house.

Empirical evidence from consumer surveys (Nielsen Media Research survey, as cited in FCC, 1998) is consistent with consumer lock-in effects. In particular, in 1997 about 80 percent of satellite subscribers rated overall satisfaction with their service as 4 or 5 out of 5. The corresponding number for cable consumers, however, was dramatically lower at about 45 percent. Yet, according to another study (Chilton Research Services Survey conducted August 11-15, 1997, as cited in FCC, 1998), only about 10 percent of cable subscribers indicated that they were "very likely" to switch to DBS. It is worth noting

[3] More details on the industry can be found in the annual Federal Communications Commission (FCC) reports.

[4] Cable service fees typically do not involve a purchase of equipment but do include its rental.

that buying services from both a satellite and a cable provider (multi-homing) is very rare in the United States because bundles are too expensive to make it reasonable.[5]

3 Model

In this section I develop a dynamic model of consumer behavior in a market with switching costs. The model allows for forward-looking consumer decisions, vertical product differentiation within service providers, strategic behavior of cable firms, and persistent consumer heterogeneity in preferences.

Flow utility

The time period, t, corresponds to one year. Every time period, each consumer can choose either cable or satellite service. The outside option is given by the *"over-the-air"* television available free of charge. Let $g \in \{o, c, s\}$ denote the outside option, cable, and satellite service providers respectively. Each of the paid-television service providers offers multiple vertically differentiated products, $j \in \mathcal{J}_g$. Products are characterized by a monthly subscription fee, p_{gjt}, and the quality of the programming content, summarized by a scalar $q_{gjt} \in \mathbb{R}^+$. In addition to the price and quality of its products, each of the television service providers is characterized by a scalar variable $\xi_{gt} \in \mathbb{R}$, representing the overall quality of its service, which is unobserved by us but observed by market participants. For example, for cable providers, ξ_{ct} may describe the quality of the customer service and availability of supplementary services, e.g., telephone and Internet. For satellite providers, ξ_{st} may represent the quality of signal reception (e.g. "southern exposure"), availability of the local TV channels, etc.

The demand side of the market is given by a continuum of heterogeneous consumers indexed by i. Let $\omega_i \overset{iid}{\sim} F_\omega$ denote a vector defining random variables describing consumer i's time-invariant preferences. Let $a_{it} \in \{o, c, s\}$ denote consumer i's choice of provider

[5] The only situations one can imagine would be in markets like Philadelphia where an American football fan would buy DirecTV to get the (exclusive) Sunday Ticket NFL package as well as a cable subscription to get the (exclusive) local Regional Sports Network.

in period t. Let $\bar{p}_{gt}$ and $\bar{q}_{gt}$ denote vectors of prices and observed quality levels for all products of a provider $g = c, s$ at time t.

Switching costs are defined similar to Klemperer (1987a, b) as constant over time provider-specific "start-up"costs, η_{ig}, known to the consumers. These costs must be paid every time the consumer chooses a new service provider. In subsequent periods, no costs are incurred if the consumer chooses any of the products from the same service provider. The decision to switch to the outside alternative is costless.

I assume that flow utility from service g is given by

$$U_i(a_{it}, a_{it-1}) = \begin{cases} -\eta_{ig} \cdot \mathbb{1}(a_{it-1} \neq g) + \tilde{\delta}_{ig}(\bar{p}_{gt}, \bar{q}_{gt}) + \xi_{gt} + \epsilon_{igt} & \text{if } a_{it} = c, s, \\ \epsilon_{iot} & \text{otherwise,} \end{cases} \tag{1}$$

where $\mathbb{1}(\cdot) \in \{0, 1\}$ is an indicator function and

$$\tilde{\delta}_{ig}(\bar{p}_{gt}, \bar{q}_{gt}) = \alpha_{ig} + \max_{j \in \mathcal{J}_{gt}} \{\alpha_{ip} p_{gjt} + \alpha_{iq} q_{gjt}\}, \tag{2}$$

with $q_{g1t} \leq q_{g2t}, \ldots, q_{g,J_{gt}-1,t} \leq q_{g,J_{gt},t}$ and $p_{g1t} \leq p_{g2t}, \ldots, p_{g,J_{gt}-1,t} \leq p_{g,J_{gt},t}$, and α_{ig} stands for the constant over time provider-specific preference parameter. This specification implies that a given consumer type i chooses one and only one product of provider g.[6] Let $\delta_{igt} \equiv \tilde{\delta}_{ig}(\bar{p}_{gt}, \bar{q}_{gt}) + \xi_{gt}$ and assume the distribution of consumer random tastes for provider, ϵ_{igt}, satisfies the following assumption.

Assumption 1: *Random shocks in the consumer utility function* (1) *are independently and identically distributed (i.i.d.) across consumers, service providers, and time and follow the Extreme Value Type 1 distribution with density* $f(\epsilon_{igt}) = \exp(-\epsilon_{igt}) \cdot \exp(-\exp(-\epsilon_{igt}))$.

Several things are worth noting. First, formulation (1) assumes additive separability of the consumer utility function in the unobservables ξ_{gt}, $g = c, s$ and $\epsilon_{it} = (\epsilon_{iot}, \epsilon_{ict}, \epsilon_{ist})$. Second, the unobserved service quality ξ_{gt} is provider-specific, i.e., it scales utility from all available service tiers and is not product-specific. This assumption is consistent with

[6] This also implies that nondegenerate product shares within the same carrier require consumer heterogeneity with respect to the price and/or quality sensitivity. This issue is discussed at length in the next section.

vertical differentiation of the services when the upper level product contains all the characteristics of the lower level ones. My interpretation of the unobserved provider-specific characteristics is related to the supplementary services offered by cable and the overall signal reception quality specific to satellite providers, which should apply to every product of a given carrier. Besides, from a practical standpoint, in order to solve for the product-specific unobservables I would need to observe shares of each tier. The latter reason is very important because I never observe tier-specific shares for satellite companies and the data on the total cable shares is more frequently available than the data on the product-specific shares. Third, a utility shock ϵ_{igt} is also provider-specific. This is also related to the vertical differentiation of the services which would not be satisfied if I allow for product-specific utility innovations. Besides, as pointed out in Ackerberg and Rysman (2005) and Berry and Pakes (2007), tastes for products represented by such shocks at the product level may have some undesirable consequences.[7]

The consumer's dynamic problem

A consumer type i maximizes the expected present discounted value of flow utilities over an infinite horizon. Let Ω_t denote current service characteristics and any other factors that affect future service characteristics. I assume that Ω_t evolves according to a first-order Markov process $P(\Omega_{t+1}|\Omega_t)$ that accounts for the providers' optimizing behavior. Let $V_i(\Omega_t, a_{it-1}, \epsilon_{it})$ denote the value function for consumer type i,

$$\begin{gathered} V_i(\Omega_t, a_{it-1}, \epsilon_{it}) \equiv \max_{\{a_{it}\}_{t=\tau}^{\infty}} \sum_{t=\tau}^{\infty} \beta^{t-\tau} \mathrm{E}\left[U_i(a_{it}, a_{it-1}) \middle| \Omega_\tau, \epsilon_{i\tau}\right] \\ \Omega_{t+1} \sim P(\cdot|\Omega_t), \quad a_{i0} = o \end{gathered} \tag{3}$$

where $U_i(a_{it}, a_{it-1})$ is defined in equation (1), $\beta \in (0,1)$ is a discount factor, and the expectation is taken with respect to future realizations of $\delta_{igt},\ g = c, s$, and ϵ_{it}.

The state vector for consumer type i is given by $(\Omega_t, a_{it-1}, \epsilon_{it})$. Then the consumer

[7] That being said, allowing the shocks to vary across products of the same provider is feasible as the model prediction of the total provider share would still be analytic under the distributional assumptions discussed below.

dynamic maximization problem (3) can be written recursively in the form of a Bellman equation,

$$V_i(\Omega_t, a_{it-1}, \epsilon_{it}) = \max_{a_{it}\in\{o,c,s\}} \{U_i(a_{it}, a_{it-1}) + \beta \mathrm{E}\left[V_i(\Omega_{t+1}, a_{it}, \epsilon_{it+1})|\Omega_t, a_{it}, \epsilon_{it}\right]\}, \quad (4)$$

where the conditional expectation is taken over the future values of state variables.

The large dimensionality of Ω_t makes the dynamic programming problem computationally intractable. To reduce the number of state variables I follow an approach similar to the one used in Melnikov (2013) and Hendel and Nevo (2006). In particular, I proceed with a simplifying assumption about the marginal distribution of state variables.

Assumption 2: *The current values of the mean utility from the two service providers are a sufficient statistic for the distribution of their future values, i.e.,* $P(\delta_{igt}|\Omega_t) = P(\delta_{igt}|\Omega'_t),\ \text{if } \delta_{igt}(\Omega_t) = \delta_{igt}(\Omega'_t),\ g = c, s.$

Assumption 2 implies that a pair of current flow utility values provides all the relevant information about the marginal distributions of future flow utility for each of the carriers. In other words, $P(\delta_{igt}|\delta_{ict}, \delta_{ist}, \Omega_t) = P(\delta_{igt}|\delta_{ict}, \delta_{ist}),\ g = c, s$. It is worth noting that this assumption is less restrictive than the so-called "dynamic inclusive value sufficiency" assumption frequently made in the literature on dynamic demand estimation (e.g., Schiraldi, 2011, Gowrisankaran and Rysman, 2012, Nosal, 2012). Contrary to the dynamic inclusive values, $(\delta_{ict}, \delta_{ist})$ summarize only prices and product characteristics (separately for each provider) and not the endogenously determined future behavior.[8]

Under Assumption 2 the state space of the problem simplifies to $(\delta_{ict}, \delta_{ist}, a_{it-1}, \epsilon_{it})$, where ϵ_{it} is a three-dimensional vector of i.i.d. innovations to the current period utility. Obviously, the assumption is fairly restrictive, as there might be other information available to the consumers about the probability distribution over future utility from paid-television services.[9] Unfortunately, incorporating such information into the model would increase the state space dramatically, making the problem computationally infeasible.

To further simplify the problem, let $V_i^g(\delta_{ict}, \delta_{ist})$ denote consumer value net of idiosyn-

[8]More details can be found in Aguirregabiria and Nevo (2013).

[9]One example of such information is government regulation of the television industry.

cratic shock ϵ_{it} when $a_{it} = a_{it-1} = g$. Using the facts that (1) switching costs depend only on the identity of the provider chosen in the current period, and (2) the properties of the i.i.d. extreme value errors, I can define a joint contraction mapping in terms of choice-specific value functions,

$$
\begin{aligned}
V_i^o &= \beta \mathrm{E} \ln \left[\exp(V_i^o) + \exp(V_i^c - \eta_{ic}) + \exp(V_i^s - \eta_{is}) \right], \\
V_i^c &= \delta_{ict} + \beta \mathrm{E} \ln \left[\exp(V_i^o) + \exp(V_i^c) + \exp(V_i^s - \eta_{is}) \right], \\
V_i^s &= \delta_{ist} + \beta \mathrm{E} \ln \left[\exp(V_i^o) + \exp(V_i^c - \eta_{ic}) + \exp(V_i^s) \right],
\end{aligned} \tag{5}
$$

where the expectation is with respect to future values of state variables $(\delta_{ict}, \delta_{ist})$ (suppressed). Note that the original value function from equation (4) now can be expressed in terms of the choice-specific value functions as follows,

$$
V_i(\delta_{ict}, \delta_{ist}, a_{it-1}, \epsilon_{it}) = \max \left\{ \begin{array}{r} -\eta_{ic} \cdot \mathbb{1}(a_{it-1} \neq c) + V^c(\delta_{ict}, \delta_{ist}) + \epsilon_{ict}, \\ -\eta_{is} \cdot \mathbb{1}(a_{it-1} \neq s) + V^s(\delta_{ict}, \delta_{ist}) + \epsilon_{ist}, \\ V^o(\delta_{ict}, \delta_{ist}) + \epsilon_{iot} \end{array} \right\}.
$$

To specify the law of motion, I assume that consumer i perceives the actual empirical marginal density of the cable and satellite flow utility δ_{ict+1} and δ_{ist+1} as a pair of simple autoregressive specifications,

$$
\delta_{ict+1} = \gamma_{0ci} + \gamma_{1ci}\delta_{ict} + \gamma_{2si}\delta_{ist} + \nu_{ict}, \tag{6}
$$

$$
\delta_{ist+1} = \gamma_{0si} + \gamma_{1si}\delta_{ist} + \nu_{ist}, \tag{7}
$$

where ν_{ict} and ν_{ist} are independently normally distributed with means 0 and variances σ_{ic}^2 and σ_{is}^2 respectively. Note that, due to the difference in strategic behavior of satellite and cable operators (the former sets prices and qualities at the national level, whereas the latter sets them at the local level), the probability distribution over the future utility flow for a cable provider is assumed to depend on both the current period own flow utility as well as the one by a satellite competitor. Hence, transition parameters are summarized by a vector $(\gamma_{0ci}, \gamma_{1ci}, \gamma_{2ci}, \sigma_{ic}; \gamma_{0si}, \gamma_{1si}, \sigma_{is})$ for each consumer type i. I conduct several robustness checks of the specifications as described in Section 6 and Appendix A.4.

Purchase probabilities and market shares

In every period, each consumer type has three possible choices and can be in one of the three possible states based on the previous period decision. This defines nine possible combinations of the current and last period choices. In particular, the probability of choosing a provider g conditional on being previously subscribed to a provider k (including the outside alternative) is given by

$$\Pr(a_{it} = g, a_{it-1} = k) = \Pr\begin{pmatrix} -\eta_{ig} \cdot \mathbb{1}(g \neq k) + V_i^g + \epsilon_{igt} \geq \\ -\eta_{il} \cdot \mathbb{1}(l \neq k) + V_i^l + \epsilon_{ilt}, \forall l \neq g \end{pmatrix}$$

$$= \frac{\exp(-\eta_{ig} \cdot \mathbb{1}(g \neq k) + V_i^g)}{\exp(V_i^o) + \exp(-\eta_{ic} \cdot \mathbb{1}(c \neq k) + V_i^c) + \exp(-\eta_{is} \cdot \mathbb{1}(s \neq k) + V_i^s)},$$

where the last line follows from the distributional assumption on ϵ_{it}.

Let s_{ict-1} and s_{ist-1} denote the shares of consumer type i subscribing to cable and satellite services in period $t-1$, respectively. Then the current period predicted market shares are given by the following expression,

$$\begin{aligned} s_{igt} = \Pr(a_{it} = g, a_{it-1} = c) \cdot s_{ict-1} + \Pr(a_{it} = g, a_{it-1} = s) \cdot s_{ist-1} \\ + \Pr(a_{it} = g, a_{it-1} = o) \cdot (1 - s_{ict-1} - s_{ist-1}), \end{aligned} \tag{8}$$

where $s_{igt} = s_{gt}(\bar{p}_{ct}, \bar{q}_{ct}, \xi_{ct}, \bar{p}_{st}, \bar{q}_{st}, \xi_{st}, s_{it-1}, \omega_i)$ and $s_{it-1} = (s_{ict-1}, s_{ist-1})$ is a pair of shares of consumer type i subscribed to cable and satellite providers in the previous period.[10]

To obtain aggregate market shares I integrate over the joint distribution of consumer heterogeneity and last period subscription, $G_{\omega, s_{it-1}}$,

$$s_{gt} = \int s_{gt}(\bar{p}_{ct}, \bar{q}_{ct}, \xi_{ct}, \bar{p}_{st}, \bar{q}_{st}, \xi_{st}, s_{it-1}, \omega_i) dG_{\omega, s_{t-1}}(\omega_i, s_{it-1}|\theta). \tag{9}$$

[10]Due to the unbounded support of the extreme value distribution, each consumer type would choose every alternative with strictly positive probability.

To compute $G_{\omega,s_{t-1}}(\cdot|\theta)$ I make an assumption about the initial conditions (initial consumer subscription) in 1992 and then numerically simulate the distribution forward till 2006.

Cable and satellite companies offer multiple products or service tiers. By assumption, consumers can switch products of the same provider at zero costs. In the data, I observe shares of individual products for cable companies only. These shares should be informative about the distribution of consumer heterogeneity with respect to price and/or quality sensitivity.

The maximum utility attained by subscribing to a provider g is defined in equation (2). Let $s_{j|g,t}$ denote the conditional market share of product j for provider g, with $\sum_{j\in\mathcal{J}_{gt}} s_{j|g,t} = 1$ by construction. Then the model's prediction for this share is described by

$$s_{j|g,t} = \frac{\displaystyle\int \mathbb{1}\left(j = \underset{j'\in\mathcal{J}_{gt}}{\arg\max}\{\alpha_{ip}p_{gj't} + \alpha_{iq}q_{gj't}\}\right) s_{igt}\, dG_\omega(\omega_i, s_{it-1}|\theta)}{\displaystyle\int s_{igt},\ dG_\omega(\omega_i, s_{it-1}|\theta)}, \tag{10}$$

where s_{igt} is given by equation (8). In estimation, I refer to the moment conditions based on the mismatch between observed and predicted product-specific shares as "heterogeneity moments."

Static versus dynamic/myopic model

Before discussing the estimation algorithm, it is worth emphasizing an important property of the proposed framework. When switching costs are zero, the model becomes identical to a static one. Let $\eta_{ic} = \eta_{is} = 0\ \forall i$ and consider the probability that consumer type i chooses provider $g \in \{o, c, s\}$ given the previous period choice of provider $k \in \{o, c, s\}$.

$$\begin{aligned}
\Pr(a_{it} = g, a_{it-1} = k) &= \frac{\mathrm{e}^{\delta_{igt}+\beta EV_i^g}}{\mathrm{e}^{\beta EV_i^o} + \mathrm{e}^{\delta_{ict}+\beta EV_i^c} + \mathrm{e}^{\delta_{ist}+\beta EV_i^s}} \\
&= \frac{\mathrm{e}^{\delta_{igt}}\,\mathrm{e}^{\beta EV_i}}{(1 + \mathrm{e}^{\delta_{ict}} + \mathrm{e}^{\delta_{ist}})\,\mathrm{e}^{\beta EV_i}} \\
&= \frac{\mathrm{e}^{\delta_{igt}}}{1 + \mathrm{e}^{\delta_{ict}} + \mathrm{e}^{\delta_{ist}}} \\
&= \Pr(a_{it} = g),
\end{aligned}$$

where the second line follows from the fact that $EV_i^o = EV_i^c = EV_i^s$ as is obvious from equation (5). In this case, the dynamic model would result in purchase probabilities for individual i and market shares defined exactly as in the static mixed (random coefficient) logit model.[11] This property of the model, where the static model is nested within a more general dynamic setup, allows me to directly test static assumptions independently of the assumptions on the discount factor. A test for myopic versus dynamic consumer behavior is much harder as it boils down to identification of the discount factor. Problems with identification of discount factors have been discussed in the literature by several authors including Manski (1993), Rust (1994), and Magnac and Thesmar (2002).

4 Data

Data for this study were compiled from several sources. The most important source is *Warren's Cable Factbooks*, which contains exhaustive information on cable systems for 1992-2006, including market size, number of subscribers, prices, channel lineups, total capacity, and identity of the owner. Data for satellite providers were collected from the Internet.[12] Data for all monetary variables were adjusted using the U.S. consumer price index with 1997 as the base year. Hence, any monetary equivalents computed in this article are in 1997 prices.

A market n is defined as the area (number of homes passed) franchised to a cable provider, and market shares for the cable firms, s_{cnt}, are computed as the ratio of cable subscribers to the total number of houses in the area. Product-specific cable shares (e.g., for Basic package, Expanded Basic-1, Expanded Basic-2) are obtained in the same way. Data on satellite penetration rates are available only at a more coarse Designated Market Area (DMA) level, which typically spans multiple markets. Therefore, to calculate satellite market shares in each market I follow Chu (2007) and assume that within a DMA satellite subscribers constitute a constant proportion of the non-cable subscribers. Thus, satellite

[11] From equation (8) it is clear that the current period share is no longer a function of the last period choices, i.e., $s_{igt} = s_{gt}(\bar{p}_{ct}, \bar{q}_{ct}, \xi_{ct}, \bar{p}_{st}, \bar{q}_{st}, \xi_{st}, \omega_i)$.

[12] I am grateful to Gregory Crawford and Ali Yurukoglu who generously shared their data with me.

market shares in market n at time t are given by

$$s_{snt} = (1 - s_{cnt})R_{dt},$$

where $R_{dt} = \frac{satsubs_{dt}}{M_{dt} - cabsubs_{dt}}$, d denotes the corresponding DMA, $satsubs_{dt}$ is the number of satellite subscribers, M_{dt} is the total number of houses, and $cabsubs_{dt}$ is the number of cable subscribers.[13] In 1997-2006 most of the markets were served by the local cable service provider and two DBS systems (DirecTV and DISH). However, in the data I observe only total share of satellite firms. Therefore, I assume that there is only one satellite competitor with the product lineup given by the products of DirecTV and total market share s_{snt} as defined above.

To reduce the computational burden I assume that observed quality is summarized by a scalar variable constructed as a weighted average of the total number of channels with the weights given by the national average fees per subscriber that television service providers must pay to carry a given channel. Channels with unknown or zero costs were assigned a cost of \$0.01. This quality measure is therefore increasing in the number of channels but also accounts for the differences in composition of the programming bundles (assuming that higher-quality channels on average cost more).

Two data issues are worth noting. First, whereas the data on aggregate market shares for cable providers are more reliable because they are reported by two distinct variables in the data sources, the data on product-specific number of subscribers contain a large number of missing observations. Second, DBS product-level shares are not available at all. Both problems are addressed in Section 5.

Descriptive statistics of the data are presented in Table 1, which reports summary statistics for updated observations used in estimation. An observation is defined as updated if the total number of subscribers is different at two consecutive time periods in 1997-2006. In total, there are 8,378 such observations. To address the initial conditions problem I used additional data from 1992-1996. These earlier data were used only to

[13] Under an alternative assumption that in all markets within the same DMA satellite market shares are the same, estimation results remain very similar to the ones reported in this article.

approximate the joint distribution of consumer types across market shares and not to form moment conditions (because of high likelihood of measurement errors in the market shares).

[*** Table 1 appears about here ***]

One important takeaway is that on average satellite and cable providers offer comparable levels of quality at relatively similar prices.[14]

Instruments

It is conceivable that local cable operators condition their policies (price and quality levels) on provider-specific service characteristics, (ξ_{cnt}, ξ_{snt}), that we cannot observe. In order to identify parameters on the endogenous price and quality variables, I use instrumental variables (IVs) similar to those suggested by Crawford (2005).[15]

The main instruments for price and quality of cable providers are average prices and quality levels of other cable systems that belong to the same multiple-system-operator (MSO). Common ownership introduces correlation in the marginal costs, and hence, in the prices and quality levels of systems within the same MSO, because the owner typically negotiates programming fees and other contract arrangements with the network providers on behalf of all of its members. To be valid instruments these variables must be uncorrelated with the unobserved local market service characteristics, ξ's. When I constructed IVs using a common ownership structure, I was unable to find other systems owned by the same MSO for about 26 percent of the systems. In these cases I used matching to construct the instruments. In particular, such systems were tabulated into 27 subgroups depending on (1) market size, (2) number of products offered, and (3) own capacity level with each of the criteria divided into small, medium, and large. Then I created the instrumental variables as if systems that belonged to the same subgroup were owned by the same MSO.

[14] After 2002, satellite providers reduced the number of packages from six to three, making programming bundles less densely located in price-quality space.

[15] The same set of instruments was used for both cable and satellite providers.

Additional instrumental variables include proxies for the bargaining power of different MSOs. Following the previous literature, I use the total number of homes passed and the number of subscribers served by the system's corporate parent as proxies for its bargaining power. Presumably, larger MSOs are able to negotiate better deals with the providers of the programming content, which should be reflected in the marginal costs of the local cable companies.[16]

MSOs with a larger average capacity level have an opportunity to get lower rates per channel because the network providers often bundle several channels. By purchasing a bundle of channels, large-capacity MSOs could reduce the costs of each channel carried by its members. I use the MSO average (across systems) capacity level as another instrumental variable.

Finally, own capacity level and the total length of own physical cable network serve as proxy variables for the differences in the maintenance costs incurred by the systems.[17] The first one is correlated with the technological sophistication of the equipment installed, whereas the second is used to control for the variation in the density of houses served by different cable providers.

To illustrate the effect the instruments have on the estimated mean price and quality sensitivity I conducted several robustness checks (OLS-like and minimal IV specifications) which are reported in Appendix A.1 and discussed in Section 6.

Identification

Identifying the parameters of consumer utility function in a dynamic model is a complex issue. As discussed in Section 3 above, in this study I don't attempt to identify the discount factor but rather assume that it is known. Another important question is the separate identification of consumer heterogeneity measured by random coefficients and switching cost parameters using only market level data. I address this issue below.

[16]For example, Crawford and Yurukoglu (2012) find that larger cable companies have about 17 percent lower input costs than their smaller competitors.

[17]Investment into fixed capital takes time and occurs infrequently because moving to fiber-optics or better compression technologies affects the entire physical network, which is a very expensive and time-consuming process.

Each consumer decision is a function of the current period payoff relevant variables and the previous period choice, i.e.,

$$\begin{aligned} a_{it} &= a\big(\delta_{ct}(x_{ct}, \xi_{ct}; \omega_i), \delta_{st}(x_{st}, \xi_{st}; \omega_i), a_{it-1}; \epsilon_{it}, \omega_i\big) \\ &= a\big(X_t, \Xi_t, H_t; \epsilon_{it}, \dots, \epsilon_{i1}, \omega_i\big), \end{aligned}$$

where $X_t = (x_{ct}, x_{st})$, $\Xi_t = (\xi_{ct}, \xi_{st})$, and $H_t = \big\{\big(X_\tau, \Xi_\tau\big)\big\}_{\tau=1}^{t-1}$ define all observed and unobserved service characteristics and the history of these characteristics starting at the initial conditions $H_1 = \varnothing$ and $a_{i0} = o\ \forall i$. Then market shares are obtained by integrating over consumer heterogeneity, i.e.,

$$s_{gt}(X_t, \Xi_t, H_t) = \int \cdots \int \mathbb{1}\Big(a\big(X_t, \Xi_t, H_t; \epsilon_{it}, \dots, \epsilon_{i1}, \omega_i\big) = g\Big) dF_\epsilon dF_\omega. \tag{11}$$

Conditional on (X_t, Ξ_t), current market shares are functions of H_t only due to switching costs. Without switching costs, persistence in consumer preferences by itself cannot account for the relationship between current choices and last period values of exogenous payoff-relevant variables conditional on their current values. If F_ϵ and F_ω are known, then the exogenous shifters of the previous period decisions are necessary for identifying switching cost parameters. In particular, I assume that the set of instruments Z_t discussed in the previous section satisfy $\mathrm{E}[\xi_{gt}|Z_t] = \mathrm{E}[\xi_{gt}|Z_{t-1}] = 0, \forall g, t$, and include moment conditions based on the lagged exogenous variables into the generalized method of moments (GMM) objective function. Intuitively, holding current prices and quality levels fixed, lower prices and higher quality levels (due to some cost reducing innovations) in the previous period would increase the previous period customer base and, hence, the current period market share. An example of the cost reducing innovations would be change in the ownership structure, which frequently occurs in the data: a previously independent local cable operator purchased by a large MSO immediately gains access to better programming at lower prices. The identification argument emphasizes the importance of the data ordering along the time dimension. Following Dube et al. (2010), I provide support to the identification strategy by estimating the model with the data randomly reordered along the time dimension (details can be found in Appendix A).

Identifying consumer heterogeneity relies on parametric restrictions. As discussed in Section 3, I assume that there is no serial correlation in consumer preferences over time (the distribution F_ϵ is i.i.d. over time and F_ω is constant). Furthermore, by assumption, the distribution of consumer heterogeneity is the same in all markets. Therefore, a standard argument that variation in observables across markets identifies parameters of persistent consumer heterogeneity applies here.[18] Finally, I include additional heterogeneity moments based on equation (13) to enhance the identification of random coefficients (see discussion in Section 5).

Reduced-form evidence of state dependence

Before discussing to the estimation, I provide reduced-form evidence of state dependence in consumer decisions which is consistent with consumer switching costs. The method extends Chamberlain's (1992) idea to the market level data, and the underlying intuition is simple.

Switching costs introduce state dependence in consumer decisions when the current period choice is a function of the previous period decision. Without state dependence, exogenous variation in the last period decisions should be irrelevant for the current period choices. Market shares represent aggregated consumer decisions. Hence, one can regress current period market shares on all contemporaneous exogenous variables and on the last period market shares instrumented with the lagged exogenous variables.[19] Statistically significant lagged market shares would suggest state dependence in consumer choices, which, in turn, may occur due to switching costs. This simple test is not consistent because it tests only for linear dependence using the following specification,

$$s_{jnt} = Z_{nt}\beta + \hat{s}_{nt-1}\alpha + \varepsilon_{jnt},$$

[18] For less parametric treatment of identification see Fox et al. (2012) and Berry and Haile (2013). It is worth noting that more flexible specifications of the unobserved persistent tastes for providers would substantially increase computational burden due to the increased number of parameters to estimate.

[19] I prefer this version of the test because it is more concise than the alternative, which would regress the current market share on the current exogenous variables and the lagged exogenous variables and test for joint significance of the latter. Results of the alternative tests are available upon request.

where Z_{nt} is a vector of exogenous variables and $\hat{s}_{nt-1}$ is a vector of market shares instrumented using Z_{nt-1}. However, if the specification passes the joint significance test for the coefficients α, this would be consistent with substantial consumer switching costs in the industry.

[*** Table 2 appears about here ***]

A summary of estimation results is presented in Table 2, with full specification and first-stage regressions available upon request.

In all specifications, lagged market shares are jointly significantly different from zero. When instrumented with the exogenous variables, lagged market shares have expected signs suggesting that, controlling for the current period covariates, larger own market share in the previous period implies larger current period own market share. Similarly, larger competitor's previous period market share ceteris paribus results in smaller own market share. Note that the critique that correlation between current choice and lagged demand shifters may simply proxy for expectations about future realizations of the payoff relevant variables does not apply here. Indeed, it is hard to see why consumers would make forward-looking decisions in the absence of any state dependence.[20]

5 Estimation algorithm

My method of estimating parameters of the model combines techniques developed by Berry et al. (1995), Rust (1987) and Gowrisankaran and Rysman (2012). Similar to Gowrisankaran and Rysman (2012), the estimation algorithm involves three levels of optimization.[21]

The inner loop solves the consumer dynamic programming problem in equation (4) for each consumer type and calculates aggregate market shares as in equation (9).[22] The

[20]Leaving any behavioral motives aside, some sort of a coordination problem may generate similar behavior. However, this is definitely not the case for the paid television industry.

[21]Note that I don't make the inclusive value sufficiency assumption but allow consumers to track the evolution of utility from each provider, separately accounting for the strategic behavior of the cable firms as discussed in Section 3

[22]Even though the algorithm always converged, as pointed out by an anonymous referee, the joint problem of solving the Bellman equation (5) and minimizing least squared errors in (6) and (7) may not have a unique solution and the GMM estimators may require a separate consistency proof.

middle loop solves for the demand side provider-specific unobservables (ξ_{ct}, ξ_{st}) that match observed market shares to the market shares predicted by the model. The outer loop searches over the parameter vector.

Below I list several additional parametric restrictions imposed in the estimation. First, instead of estimating the discount factor, β, its value is set to 0.90.[23] Second, I assume that $\eta_{ic} = \eta_c$ and $\eta_{is} = \eta_s$ $\forall i$. Because the switching cost parameters are not interacted with any of the observable variables, it is very hard to see which variation in the data could help identify second moments of their distribution. Third, to simulate consumer types, I assume that the parameters of consumer heterogeneity are represented by random draws from a distribution known up to a parameter vector. In particular, $\omega_i = (\alpha_{ic}, \alpha_{is}, \alpha_{ip}, \alpha_{iq})'$ has a multivariate normal distribution with a diagonal variance-covariance matrix Σ, i.e., $\omega_i \overset{iid}{\sim} N(\bar{\alpha}, \Sigma)$. The vector of parameters to estimate is $(\bar{\alpha}_c, \bar{\alpha}_s, \bar{\alpha}_p, \bar{\alpha}_q, \sigma_{\alpha_c}, \sigma_{\alpha_s}, \sigma_{\alpha_p}, \sigma_{\alpha_q})$.[24] In cases when individual level data are available, more flexible distributions of persistent consumer heterogeneity can be used (e.g., Dube et al., 2009, 2010). With aggregate data, functional form restrictions become more important because nonparametric identification arguments (e.g., Berry and Haile, 2013) are not directly applicable to the models with state dependence in consumer preferences. It is worth noting that identification of the unobserved price and quality sensitivity does not come solely from the functional form restrictions and is facilitated by additional moment conditions based on product-specific shares defined in equation (10).

To address the initial conditions problem in the models with unobserved persistent heterogeneity in consumer preferences, I used data from 1992-1996 and an assumption that DBS entered in 1994. Robustness checks for the alternative assumptions about the date of satellite entry can also be found in Appendix A.2.

[23] An alternative high value for the discount factor of 0.95 was tested with the results not substantially different from the ones reported in Section 6 and substantial increases in the estimation time.

[24] Off-diagonal elements of Σ turn out to be statistically insignificant; results are available from the author upon request.

Moment conditions

For a given vector of parameters, an estimation algorithm can be used to compute provider-specific characteristics (ξ_{ct}, ξ_{st}) that only market agents can observe. Similar to Berry (1994), Berry et al. (1995) and the literature that follows, I use these structural errors to generate moment conditions for estimation. In the model, I allow only for provider- and not product-specific unobservables. Hence, solving for the shocks requires matching model predictions for only the aggregate provider-specific shares to the ones observed in the data.

Let $\xi_{ct}(\theta)$ and $\xi_{st}(\theta)$ denote values of the unobservables solving the following system of equations,

$$\begin{cases} s_{ct} = s_{ct}(\bar{p}_{ct}, \bar{q}_{ct}, \xi_{ct}, \bar{p}_{st}, \bar{q}_{st}, \xi_{st}|\theta), \\ s_{st} = s_{st}(\bar{p}_{ct}, \bar{q}_{ct}, \xi_{ct}, \bar{p}_{st}, \bar{q}_{st}, \xi_{st}|\theta), \end{cases} \tag{12}$$

where s_{ct} and s_{st} are observed cable and satellite market shares and $s_{ct}(\cdot)$ and $s_{gt}(\cdot)$ are the model predictions. Using proofs similar to the ones in Berry (1994) and Berry et al. (1995), it is straightforward to show that static and myopic consumer models satisfy sufficient conditions (strict monotonicity and substitution) for a unique solution to (12). This is trivially true for the initial period, $t = 1$. By conditioning on the solution to the previous period system of market share equations, one can inductively establish the inverse for any $t > 1$. In a dynamic framework with general transition probability functions, this is less obvious due to the possibility of intertemporal substitution effects, which may violate the monotonicity assumptions. An example of sufficient conditions for $\partial s_{gt}/\partial \xi_{gt} > 0$ and $\partial s_{gt}/\partial \xi_{kt} < 0, k \neq g$ would require increasing (in the sense of first-order stochastic dominance) transition probability functions in the per-period reward function, δ_{igt}, for each consumer type i.[25]

To estimate parameters of the consumer utility function I make the following assumption about the unobservables.

[25]Parameter estimates for the reduced-form consumer beliefs (available from the author upon request) are consistent with this assumption.

Assumption 3: *For each service provider, the unobserved service characteristics can be written as* $\xi_{gnt} = \bar{\alpha}_g + \alpha_{gt} + \tilde{\xi}_{gnt},\ g = c, s,$ *where* $\bar{\alpha}_g$ *denotes provider-specific intercept,* α_{gt} *is provider-specific time effect, and* $\tilde{\xi}_{gnt}$ *satisfies mean independence assumption* $\mathrm{E}[\tilde{\xi}_{cnt}|z_{cnt}] = \mathrm{E}[\tilde{\xi}_{snt}|z_{snt}] = 0$, *for the vectors of instrumental variables* z_{cnt} *and* z_{snt}.

Note that the assumption allows not only for the provider-specific intercept but also for provider-specific time effect. The intuition is that cable and satellite technology may be quite different from the consumers' viewpoint, e.g., Internet and telephone service offered by cable but not by satellite. Time effects would account for the industry regulation (e.g., permission to broadcast local channels for DBS) development in television programming (e.g., popular TV shows and nationwide political and sport events) as well as variation in the outside alternative (e.g., development of video rental services). Importantly, these time effects can only represent shocks unanticipated by the consumers. The latter is due to the assumption on consumer beliefs (6) and (7), which do not depend explicitly on time.[26] In other words, the time effects enter the perceived transition functions of consumers (6) and (7) through the terms ν_{ict} and ν_{ist}.[27]

Because the paid television service providers offer multiple products, an additional set of moment conditions can be used in estimation. Specifically, I assume

$$s_{j|c,t} = \hat{s}_{j|c,t}(\bar{p}_{ct}, \bar{q}_{ct}, \xi_{ct}, \bar{p}_{st}, \bar{q}_{st}, \xi_{st}|\theta) + u_{cjt}, \quad s.t.\ \mathrm{E}[u_{cjt}|\mathcal{I}_t] = 0, \tag{13}$$

where $\hat{s}_{j|c,t}(\cdot)$ is the model prediction and $\mathcal{I}_t$ stands for all available information. Equation (13) defines approximation (and measurement) errors $u_{gjnt},\ g = c, s,\ j = 1, \ldots, J_{nt}$ in the predictions of individual product shares, i.e., shares of each service tier given the choice of provider. The approximation error may occur because I assume the same distribution of consumer types in every market, which helps in identifying persistent consumer heterogeneity. However, there is another important reason for the "nonstructural" interpretation of errors in the predictions of conditional shares. As discussed in Section 4, observed

[26] It might be possible to include time as an additional state variable in the consumer dynamic optimization problem. However, this would require some assumptions about the end-period values and would increase the computational burden dramatically.

[27] Ideally, one would also allow for market fixed effects. However, as discussed in Section 4, the number of reliable observations per market is very small and cross-sectional variation is important for the identification of the structural parameters.

product-specific shares for cable providers are measured with errors. Given the highly nonlinear model structure, equation (13) provides a trade-off between retrieving useful information available in the data on conditional market shares and avoiding structural interpretation of the mismeasurements.[28] Note that, according to the model structure, the only way to get nondegenerate conditional product shares is to allow for consumer heterogeneity with respect to price and/or quality sensitivity. These additional heterogeneity moments should be informative about the distribution of consumer preferences.

Let K_c, K_s, and K_u denote the number of instrumental variables for cable and satellite shocks and the number of variables to interact with the approximation error respectively. Then, $G_c(\theta) \equiv \mathrm{E}[\tilde{\xi}_{cnt}(\theta) \cdot z_{cnt}]$, $G_s(\theta) \equiv \mathrm{E}[\tilde{\xi}_{snt}(\theta) \cdot z_{snt}]$, and $G_u(\theta) \equiv \mathrm{E}[u_{cjnt}(\theta) \cdot z_{ujnt}]$ are $(K_c \times 1)$, $(K_s \times 1)$, and $(K_u \times 1)$ vectors of population moment conditions, which are assumed to be zero at the true parameter vector θ_0. Define $G^N(\theta) = \left[G^N_c(\theta), G^N_s(\theta), G^N_u(\theta)\right]'$ to be sample analog of the population moment conditions, which are then used to form the GMM objective function for the outer loop

$$\hat{\theta}^* = \arg\min_{\theta} \left\{ G^N(\theta)' \cdot W \cdot G^N(\theta) \right\}, \tag{14}$$

where W is the optimal weighting matrix.

6 Results

In this section, I focus on the most important parameters of the model.[29] First, I present the results for a one-type model. The results for static, myopic, and two versions of the dynamic model are listed in Table 3.

[*** Table 3 appears about here ***]

The results suggest that both the dynamic and myopic models predict positive switching costs for DBS of similar magnitude. In the myopic model, switching costs for cable providers

[28] Any structural tier-specific unobservable (something observed by consumers but not by us) would be severely contaminated by the measurement error.

[29] Full results for all estimated specifications are available upon request.

are not significantly different from zero, whereas in the dynamic model the estimate is positive and significant.[30] The representative consumer dynamic model predicts cable switching costs of \$159 and \$193, whereas for satellite these numbers are \$213 and \$242 for the models without and with heterogeneity moments, respectively.

Static and myopic models overestimate (in absolute value) price and quality coefficients relative to the dynamic specifications. Intuitively, this occurs because variation in the aggregate consumer decisions must be explained by the (exogenous) contemporaneous variation in the relevant variables only, i.e., potential effect of the change on the future flow utility is ignored in static and myopic specifications. When consumers make forward-looking decisions, an increase in price in the current period is informative about the future utility values. For example, according to the estimates of the reduced-form consumer beliefs (available from the author upon request), an increase in price lowers current period utility and implies lower utility in the future. Therefore, the observed responses of the market shares explained by the contemporaneous change alone within the static and myopic setup is now divided between a lower current flow utility and less optimistic expectations about the future utility.

To assess the economic significance of the parameter estimates across different models, I calculate several versions of price elasticity. Table 4 reports price elasticities evaluated for the market with identifier *IA0192* in year 2000 with the following characteristics: $s_c = 0.51,\ s_s = 0.21,\ \hat{\delta}_c = 0.60,\ \hat{\delta}_s = 0.28,\ p_c = 27.96,\ q_c = 3.50$.[31] The system belongs to the MSO "Mediacom" and is a typical (average) observation in my sample.

[*** Table 4 appears about here ***]

The estimates suggest that both the static and myopic specifications overestimate short-run price elasticity. The long-run elasticity for dynamic specification is larger than the short-run elasticity because it accounts for the effect of the current period price change on the continuation value in the consumer dynamic programming problem. In general, the difference between the short-run and long-run elasticity in a dynamic model depends

[30]When heterogeneity moments are included, standard errors become much smaller.

[31]The system is located in Iowa, DMA Quincy-Hannibal-Keokuk, community Montrose, ZIP code 52639.

on the consumer beliefs about future flow utility (as a function of the current period utility). For example, if lower current period utility implies lower expected future utility, a consumer facing a price increase (lower utility) today would also expect lower utility in the future. Therefore, when the price increase is permanent it has a larger effect on the consumer's decisions, which is reflected in the estimates of short- and long-run elasticity.

Appendix A provides robustness analysis using a representative consumer model. In particular, I illustrate the effect of different sets of instrumental variables on the parameter estimates (Table 9), provide evidence of the importance of lagged exogenous variables for identification of the switching costs (Table 11), and test alternative specifications of consumer beliefs in equations (6) and (7) (Tables 12 and 13).

[*** Table 5 appears about here ***]

Table 5 summarizes estimates for various versions of consumer heterogeneity. In alternative specifications I attempted to estimate standard deviations of the random coefficients on the provider-specific constant terms but they turn out to be not statistically significant. Estimates of the switching cost parameters become slightly bigger, but the difference is not sizable. The parameters are generally estimated accurately. Including the heterogeneity moments makes parameter estimates substantially more precise in terms of smaller standard errors. What is less encouraging are the values of the Hansen J-statistic, which become so large that none of the specifications passes the overidentifying restrictions test. Most likely the assumption that consumer heterogeneity is the same across all markets is too strong. One way to relax this assumption would be to group markets according to some observable demographic characteristics and let parameters of consumer heterogeneity vary across (but not within) these groups. Although more flexible, this approach would require solving a separate Bellman equation not only for each consumer type but also for each type in each group of the markets, which would substantially increase the computational burden. Unfortunately, the available data are unlikely to be sufficient to render more flexible estimation.

An important message from the estimation results is that the estimates of consumer switching costs are robust to various forms of heterogeneity in consumer preferences. Table

6 summarizes estimates of consumer switching costs from all specifications.

[*** Table 6 appears about here ***]

One way to assess the magnitudes of switching costs for cable and satellite is to relate them to the cost of a one-year subscription. For example, a consumer subscribed to a typical Expanded Basic cable service with a monthly fee of \$28.05 would pay about \$336.60 a year. In this case, the cost of switching to the cable service is equivalent to about half a year in terms of variable cost of service. Similarly, for a DBS subscriber to the package "Total Choice" with price of \$31.99, annual spending adds up to \$383.88, so that a one-time switch to DBS is also about one-half of the annual service cost.

7 Counterfactual simulations

To evaluate the effect of consumer switching costs on the market outcomes, I develop a simple dynamic supply side model assuming a representative consumer demand model. The model is fairly stylized and cannot fully account for the changes in the cost structure across alternative scenarios, i.e., it should be treated as a partial equilibrium analysis. However, the model is dynamic and is designed to illustrate important implications of consumer switching costs for the optimal policy of cable service providers.

Under a set of assumptions, detailed in the Appendix B, the supply side dynamic programming problem can be expressed as

$$W(s_{ct-1}, s_{st-1}, \delta_{st}, \Psi_{ct}) = \max_{\delta_{ct}} \left\{ \begin{array}{l} s_c(s_{ct-1}, s_{st-1}, \delta_{st}, \delta_{ct}) \left(\dfrac{1}{\alpha_p} \delta_{ct} + \Psi_{ct} \right) \\ +\beta \, \mathrm{E}\big[W(s_{ct}, s_{st}, \delta_{st+1}, \Psi_{ct+1} | \delta_{st}, \Psi_{ct}\big] \end{array} \right\}.$$

with the state variables given by lagged market shares s_{ct-1}, s_{st-1}, mean utility from satellite service, δ_{st}, and a variable Ψ_{ct} defined as "cost structure."[32] Mean per-period consumer utility from cable service, δ_{ct}, becomes a dynamic control.

[32] A much more complicated framework would be required to account for persistent consumer heterogeneity (Shcherbakov, 2009).

In general, Ψ_{ct} is a function of exogenous cost shifters Z_{ct} and unobserved service characteristics ξ_{ct}. It also encapsulates optimal quality choice, which is itself a function of (Z_{ct}, ξ_{ct}). By making a parametric assumption on $P(\Psi_{ct+1}|\Psi_{ct})$ it is possible to solve for $\widehat{\Psi}_{cnt}$ that rationalizes the "observed" choice of δ_{cnt} in the sample. All counterfactual simulations discussed below are conducted under the assumption that the evolution of exogenous variables and, hence, of Ψ_{ct} is constant across scenarios.

The supply side estimates are then used to conduct counterfactual simulations addressing several economic questions, including the effects of switching costs and DBS entry on cable prices and consumer utility. To disentangle these effects, I simulate four counterfactual scenarios representing hypothetical monopoly and duopoly market structures. Each of the scenarios is evaluated at both the observed (short-run) and steady state (long-run) industry outcomes. Simulating industry counterfactuals in steady state is motivated by the fact that DBS providers entered the industry relatively recently (at least for the data set used for estimation). Therefore, the market structure under consideration changed substantially when the share of satellite companies grew over time.[33] The long-run values of the exogenous variables Ψ_{ct} and δ_{st} are obtained from the reduced-form consumer beliefs. Note that the long-run simulation results can only be expressed in terms of alternative values for the consumer mean utility, δ_{ct}, because at the steady-state value of Ψ_{ct} new optimal quality levels are not separately identifiable from Z_{ct} and ξ_{ct}. In the short run, at the observed values of Ψ_{ct} quality stays constant across scenarios because (Z_{ct}, ξ_{ct}) are assumed to be constant. Therefore, for the same values of Ψ_{ct} any change in δ_{ct} comes solely from the change in price with the scale factor of α_p.

Long-run analysis

The first scenario, labeled DD, corresponds to a dynamic duopoly. Under DD, consumers have to pay switching costs when changing service providers, cable faces competition from a (nonstrategic) DBS provider, and variables δ_{st} and Ψ_{ct} follow first-order Markov

[33] At steady state, values of the "exogenous" variables $\delta_s^* = 1.218$ and $\Psi_c^* = 2.121$ cable and satellite market shares would be 0.42 and 0.45, respectively. This is in contrast to the market shares observed in the data with average $s_{ct} = 0.51$ and $s_{st} = 0.169$.

processes with the estimated coefficients.

The second scenario is labeled DM. It corresponds to dynamic monopoly, where consumers still have switching costs, but there is no satellite competitor. This simulation evaluates the effect of satellite entry on cable policy. Without DBS cable firms would obviously change their pricing policy. Because consumers are making forward-looking decisions they would take that into account by updating beliefs about future flow utilities accordingly. Therefore, to find an equilibrium optimal policy for a dynamic monopolist I first compute it holding consumer beliefs fixed and then update consumer beliefs given the new DM policy. These iterations are repeated until reduced-form consumer beliefs and optimal producer policy are consistent with each other.

The third scenario is labeled SM for static monopoly. It shows the effect of switching costs on the optimal pricing of a monopolist. In particular, I compare policies for a dynamic monopolist, DM, with the one for a static monopolist, SM. Under the latter scenario, consumers do not have switching costs, and both the monopolist and consumers solve simple static maximization problems.

The least realistic situation is when consumer switching costs are set to zero resulting in a static optimization problem on the demand side, whereas both DBS and cable companies are still competing on the supply side. This scenario is labeled as SD referring to a static duopoly. Without information about cost structure of the satellite provider it is impossible to find a new optimal DBS policy. Hence, I proceed under assumption that the DBS optimal policy δ_{st} remains the same as under DD.

In equilibrium without a DBS competitor, single-product cable firms would offer 51 percent lower mean utility than under dynamic duopoly. Hence, satellite entry indeed imposed significant competitive pressure on the incumbent natural monopolists.

By comparing equilibrium policies under dynamic and static monopoly, one can conclude that utility generated by a dynamic monopolist would exceed the one offered by its static counterpart by as much as 89 percent. In other words, the presence of consumer switching costs in a monopoly market dramatically increases optimal mean utility from the service through lower prices. Intuitively, to attract consumers facing switching costs a monopolist must heavily subsidize them by offering lower service fees.

Finally, by comparing dynamic and static duopoly (assuming that satellite policy does not change), I conclude that consumer switching costs increase consumer mean utility (due to lower prices) by about 40 percent. Simulation results are summarized in Table 7.

[*** Table 7 appears about here ***]

If one is willing to assume that the steady state value of $\Psi_c^* = 2.121$ is sufficiently close to the current sample average of 2.183, so that the equilibrium quality level of single-product cable providers can be fixed at $q_c = 3.72$ and $\xi_c = 0.67$ rationalizing price of \$24.15 at the average mean flow utility of $\delta_c = 0.431$ for this group of firms, then the differences in optimal mean utility can be translated into the differences in prices. In particular, in the steady state of scenario DD, cable price will be lower by 37 percent relative to the observed one, i.e., \$15.20 instead of \$24.15. Under static duopoly the cable firm would charge \$21.03 or 38 percent more. In markets without the DBS competitor, the cable firm would charge \$22.56, which is 48 percent higher than the predicted dynamic duopoly outcome. The effect of switching costs on a single-product cable firm turned out to be much smaller in prices than in utility terms, with an unregulated static monopolist charging \$28.91 instead of \$22.56.

Short-run analysis

Another possible simulation could evaluate optimal cable policy under different scenarios at the values of $(s_{ct}, s_{st}, \delta_{st}, \Psi_{ct})$, which are actually observed in the data. The inferences from such an exercise may be less meaningful than the steady-state analysis above because the observed duopoly share of cable providers may be fairly far from the one predicted in the steady state, which would be true for the other variables as well. However, this exercise may be quite useful to illustrate the importance of dynamics in the consumer and producer decisions.

Consider the following three scenarios: (1) DBS exits the market and all its consumers take the outside option, i.e., the cable market share stays at its observed level, (2) DBS stays in the market but consumer switching costs are completely subsidized, and (3) DBS exits the market and cable switching costs are subsidized. By analogy to the steady-state counterfactuals, I label case (1) with DM^o, case (2) with SD^o, and case (3) with SM^o. Consumer beliefs in the dynamic monopoly scenario are updated to be consistent with the optimal policy of the cable monopolist. Sample average of the observed (dynamic duopoly) policy is denoted as DD^o. Table 8 summarizes the results.

[*** Table 8 appears about here ***]

The importance of dynamics can be seen from the line comparing dynamic monopoly versus observed outcome. By assumption the difference between DD^o and DM^o is due to the removal of DBS from the market and attributing its share to the outside option. Low cable share (relative to the optimal for the monopoly scenario) creates a very strong incentive to invest into the customer base. Therefore, it would price substantially below what is currently observed. However, at some point this incentive would be offset by the "harvesting" one and the policies will be consistent with what is reported in Table 7. Because the experiment was conducted for the values of δ_s, which are far from the expected steady state, the effect of switching costs appears quite small. However, in the long-run as $\delta_s \to \delta_s^*$, market shares distribution and the resulting optimal policies for static and dynamic duopoly would diverge further apart to the values suggested by Table 7.

To see how the differences in the demand side parameter estimates translate into the differences in the counterfactual conclusions, I recover the underlying cost structure using parameter estimates for the static demand side model (listed in column 2 of Table 3). Then I compute counterfactual static duopoly and monopoly policy. Under the duopoly, the long-run cable and satellite shares would be 0.30 and 0.63, respectively. Under monopoly the cable share would be 0.60 and the equilibrium mean utility level would be 73 percent lower (against 51 percent difference predicted for the market with switching costs). Hence, a researcher ignoring consumer switching costs would significantly overestimate the effect of DBS entry on cable prices.

Finally, by using simulated optimal policy for static and dynamic monopoly I compare them as a function of the customer base. As one would expect, in comparison to a static monopolist, a dynamically optimizing monopolist may offer lower or higher prices (higher or lower utility, respectively) depending on the size of its market share. Such behavior is consistent with the investment and rent harvesting incentives described by the theoretical literature on switching costs (see Appendix B, Figure 1).

8 Conclusions

I develop and estimate a dynamic model of consumer behavior in markets with consumer switching costs. Using the estimated structural demand side parameters and a simple dynamic supply side model I conduct several counterfactual simulations evaluating the effects of switching costs and DBS entry on the equilibrium cable policies.

The estimates suggest that consumer switching costs are economically and statistically significant and constitute roughly half of the annual variable service costs for each of the service providers. Professional installation fees in the cable industry amount to only about 20 percent of the total monetary value of switching costs, with the remaining part representing unobserved hassle utility costs.

Counterfactual results based on a simplified supply side model suggest that the magnitude of the switching costs I estimate could be significant in terms of optimal price and quality choice for cable firms. In particular, due to switching costs cable firms have an incentive to invest heavily into the customer base, which may increase consumer utility from the cable service by almost 40 percent relative to the no-switching-costs situation. The investment incentive is found to be comparable to the effect of satellite entry, which induces cable firms to provide almost 50 percent higher consumer utility relative to a hypothetical monopoly situation.

It is worth noting that the consumer dynamic problem developed in this article is not capable to forecast the most recent technological innovations in the paid-television industry, including Internet television, video on demand, and pay-per-view interactive systems. Therefore, findings in this study should be applied with caution to the industry evolution beyond 2006. Finally, the role of bundling television, telephone, and Internet services by the cable firms, currently modeled in a reduced form as unobserved provider characteristics, is an interesting question left for future research.

Table 1: Summary statistics, 1992-2006.

variable	**min**	**max**	**mean**	**sd**
# of time periods	7	15	12.15	2.34
# of cable products	1	3	1.43	0.62
# of satellite products	3	6	5.38	1.21
cable market share	0.051	0.940	0.556	0.185
satellite market share	0.0001	0.693	0.109	0.106
cable price tier 0	2.68	121.30	18.58	6.63
cable price tier 1	4.76	116.51	28.05	11.66
cable price tier 2	6.33	291.08	52.13	23.25
satellite price tier 0	14.44	39.244	20.15	8.02
satellite price tier 1	19.26	43.61	25.71	7.84
satellite price tier 2	28.89	87.22	42.27	20.05
cable quality tier 0	0.010	13.130	2.659	1.654
cable quality tier 1	0.123	15.549	4.930	2.222
cable quality tier 2	2.097	17.562	8.619	3.260
satellite quality tier 0	1.782	11.727	4.520	3.197
satellite quality tier 1	3.3	12.674	5.844	2.992
satellite quality tier 2	5.65	27.882	9.545	5.714
own miles of cables ('000)	0.0003	6.802	0.117	0.294
own capacity	8	415	40.978	16.213
before 2002				
satellite price tier 3	32.75	43.80	36.45	3.74
satellite price tier 4	38.53	46.59	41.84	2.603
satellite price tier 5	46.23	77.35	54.72	11.382
satellite quality tier 3	5.93	10.442	7.235	1.674
satellite quality tier 4	6.53	10.965	7.835	1.641
satellite quality tier 5	6.96	15.436	9.128	3.280

Notes: In total, the data contains information on 2716 markets observed from 1992 till the last year with the updated observation in a given market. For example, suppose in a given market updated observations are found in 1999, 2000, 2002, and 2003. Then, this market is assumed to be observed from 1992 till 2003 and would contribute to four individual moment conditions in total. Individual moment conditions for 1997, 1998, and 2001 are ignored for this market because of the measurement error problem. The remaining eight observations (1992-1998 and 2001) are used only to approximate the initial conditions for the updated observations. In summary, in this market I would use 12 observations to keep track of the joint distribution of consumer types and market shares (from 1992 till 2003) with only four of them included as sample moments.

Table 2: Reduced-form evidence of state dependence in aggregate consumer decisions, 1992-2006.

	OLS in (1) – (4)				IV in (5) and (6)	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	s_{ct}	s_{st}	s_{ct}	s_{st}	s_{ct}	s_{st}
s_{ct-1} or $\hat{s}_{ct-1}$	0.931***	-0.0232***	0.740***	-0.0221***	0.789***	-0.195***
	(0.00266)	(0.00122)	(0.00553)	(0.00257)	(0.129)	(0.0575)
s_{st-1} or $\hat{s}_{st-1}$	0.0325***	0.994***	0.105***	0.870***	0.119	0.875***
	(0.00863)	(0.00394)	(0.0111)	(0.00514)	(0.220)	(0.105)
Current period Z_t	Yes	Yes	Yes	Yes	Yes	Yes
Time effect	Yes	Yes	Yes	Yes	Yes	Yes
Market FE	No	No	Yes	Yes	Yes	Yes
Observations	28,548	28,548	28,548	28,548	28,548	28,548
R-squared	0.866	0.915	0.515	0.895	0.513	0.873
F-test	84,753.46	48,474.26	9,979.39	17,485.53	122.47	482.86
(p-val)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
No of markets			2,716	2,716	2,716	2,716

Standard errors in parentheses: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3: Estimation results for a representative consumer model. Two-step optimal GMM, second stage.

	One-consumer-type models			
Variable	static	myopic	dynamic (1)	dynamic (2)
switching cost cab		-0.408	1.031	1.177
		(0.470)	(0.440)	(0.208)
switching cost sat		1.686	1.384	1.470
		0.411	(0.280)	(0.210)
constant cab	0.825	0.707	0.615	0.506
	(0.162)	(0.200)	(0.162)	(0.055)
constant sat	-1.026	0.299	-0.719	-0.993
	(0.540)	(0.630)	(0.400)	(0.152)
price	-0.114	-0.119	-0.078	-0.073
	(0.005)	(0.011)	(0.010)	(0.002)
quality	0.378	0.389	0.277	0.273
	(0.016)	(0.030)	(0.025)	(0.009)
time effects	yes	yes	yes	yes
heterogeneity moments	no	no	no	yes
cab switching costs, USD'97		-41.14	158.62	193.48
sat switching costs, USD'97		170.02	212.92	241.64
Hansen J-stat	71.363	52.893	38.976	96.061
p-val	0.0008	0.0344	0.3374	0.0003

Table 4: Cable and satellite own price elasticities.

variable	type	estimate	
		cable (s.e.)	satellite (s.e.)
E^{STATIC}	static	-1.57 (0.07)	-1.83 (0.08)
E^{MYOPIC}	myopic	-1.44 (0.13)	-1.94 (0.18)
$E^{DYN\ SR}$	dynamic short-run	-0.95 (0.03)	-1.13 (0.03)
$E^{DYN\ LR}$	dynamic long-run	-1.14 (0.03)	-2.32 (0.06)

Notes: Static price elasticity, E^{STATIC}, is calculated using a price coefficient estimated under the assumption of no switching costs. Myopic price elasticity, E^{MYOPIC}, is obtained using a price coefficient from a model where $\beta = 0$, i.e., consumers disregard future effects of their current period decisions. Dynamic short-run elasticity, $E^{\mathrm{DYN\ SR}}$, is computed under the assumption that the increase in price is unanticipated and nontransitory. Finally, dynamic long-run elasticity, $E^{\mathrm{DYN\ LR}}$, assumes unanticipated permanent price increase. Formulas for each type of cable price elasticity are available from the author upon request.

Table 5: Multi-type consumer model, two-step optimal GMM, 2nd stage.

Variable	Heterogeneous consumer types					
	(1)	(2)	(3)	(4)	(5)	(6)
switching cost cab	1.220	1.151	1.196	1.339	1.370	1.389
	(0.431)	(0.448)	(0.443)	(0.238)	(0.226)	(0.302)
switching cost sat	1.506	1.392	1.487	1.584	1.392	1.610
	(0.318)	(0.292)	(0.323)	(0.273)	(0.238)	(0.326)
constant cab	0.611	0.601	0.619	0.565	0.460	0.539
	(0.178)	(0.170)	(0.181)	(0.090)	(0.076)	(0.073)
constant sat	-0.986	-0.893	-0.956	-1.205	-1.400	-1.325
	(0.393)	(0.400)	(0.398)	(0.309)	(0.310)	(0.179)
price	-0.074	-0.076	-0.075	-0.071	-0.068	-0.070
	(0.009)	(0.009)	(0.009)	(0.002)	(0.002)	(0.004)
std(price)	0.012		0.012	0.016		0.016
	(0.004)		(0.004)	(0.001)		(0.002)
quality	0.273	0.273	0.274	0.271	0.269	0.272
	(0.020)	(0.023)	(0.021)	(0.011)	(0.011)	(0.013)
std(quality)		0.028	0.00001		0.037	0.019
		(0.016)	(0.650)		(0.002)	(0.004)
time effects	yes	yes	yes	yes	yes	yes
heterogeneity moments	no	no	no	yes	yes	yes
cab switching costs, USD'97	197.84	181.74	191.36	226.31	241.76	238.11
sat switching costs, USD'97	244.22	219.79	237.92	267.72	245.65	276.00
Hansen J-stat	38.06	37.50	38.28	94.91	98.12	96.49
p-val	0.332	0.355	0.281	0.0003	0.0001	0.0001

Table 6: Monetary value of switching costs for alternative specifications

model	htg-moments	cable	satellite
Representative consumer	no	158.62	212.92
	yes	193.48	241.64
Random coef. on price	no	197.84	244.22
	yes	226.31	267.72
Random coef. on quality	no	181.74	219.79
	yes	241.76	245.65
Random coef. on price & quality	no	191.36	237.92
	yes	238.11	276.00

Table 7: Optimal policy of cable provider and equilibrium market shares under alternative scenarios, evaluated at $\delta_s^* = 1.218$ and $\Psi_c^* = 2.121$

scenario	cab transition	sat transition	steady state values		
			s_c^*	s_s^*	δ_c^*
Dynamic Duopoly, DD	$\gamma_{0c} = 0.176$ $\gamma_{1c} = 0.792$ $\gamma_{2c} = 0.059$ $\sigma_c = 0.363$	$\gamma_{0s} = 0.189$ $\gamma_{1s} = 0.845$ $\sigma_s = 0.217$	0.42	0.45	1.129
Static Duopoly, SD			0.31	0.53	0.674
Dynamic Monopoly, DM	$\gamma_{0c} = 0.136$ $\gamma_{1c} = 0.762$ $\sigma_c = 0.465$		0.69	0	0.555
Static Monopoly, SM			0.52	0	0.060
DM vs DD	DM offers 51% lower utility (48% higher prices)				
SM vs DM	SM offers 89% lower utility (28% higher prices)				
SD vs DD	SD offers 40% lower utility (38% higher prices)				

Table 8: Optimal policy of cable provider under alternative scenarios evaluated at sample averages: $s_c = 0.51, s_s = 0.17, \delta_s = -0.034, \Psi_c = 2.18$

scenario	δ_c^*	comparison
Sample average, DD^o	0.431	observed outcome
Static duopoly, SD^o	0.414	mean utility is about 4% lower than DD^o
Dynamic monopoly, DM^o	0.695	mean utility is about 61% higher than observed
Static monopoly, SM^o	0.089	mean utility is about 79% lower than observed

Appendix A Robustness analysis

In this appendix I discuss several robustness checks conducted using a representative consumer dynamic model. First, I illustrate the effect of the instrumental variables used in estimation. Second, I test specifications of consumer beliefs in equations (6) and (7) at the estimated parameter values. Finally, the importance of the lagged exogenous instruments for the identification of switching costs is explored.

The effect of the instruments on the parameter estimates is illustrated in Appendix A.1 in Table 9, where I report OLS-type estimation results and another specification with a smaller number of instrumental variables (only average MSO price and quality levels and their lags were used). The comparison was conducted using a representative consumer model with heterogeneity moments included. According to the results, the instrumental variables work as expected. When a smaller number of IVs is used parameter estimates do not change much but standard errors become larger.

Specifications of consumer beliefs in equations (6) and (7) are fairly restrictive. To test whether future utility flows can be predicted significantly better using additional covariates or more lags, I use the recovered equilibrium sequences of per-period utility flows $(\delta_{cnt}, \delta_{snt})$. Results of the exercise can be found in Appendix A.4 in Table 12 and Table 13. In particular, Table Appendix A.4 presents results on various alternative specifications that additionally include price and quality levels as explanatory variables. Table 13 reports specifications with higher order lags of the cable and satellite flow utility.

Overall, the results suggest that additional covariates and higher order lagged values of the flow utility may have statistically significant effects on the current period flow utility in addition to the explanatory variables already included in equations (7) and (8). At the same time, the model fit as measured by R^2 does not increase substantially. For example, for the equation describing cable utility evolution original R^2 of 0.662 increases to 0.672 when cable prices for each tier are included as extra regressors, and further to 0.679 when adding observed satellite prices and quality levels. For the equation approximating the evolution of satellite utility, the original specification has R^2 of 0.842. By adding lagged average (across all providers) cable flow we improve its fit to 0.843. Including lagged satellite price and quality levels increases R^2 to 0.845. Finally, when both cable and satellite price and quality levels are included in addition to the covariates specified by equation (8) the fit improves to 0.848. Apparently, additional covariates and extra lags marginally improve predictions of the future utility flows. At the same time, adding extra variables into the state space would make numerical solution to the dynamic programming problem much harder.[34] In addition, extra lags would reduce the number of observations used to estimate the model considerably given that I already have a very short panel.

Discussion in Section 4 emphasizes the importance of the moment conditions based on the lagged exogenous variables for the identification of switching costs. To illustrate the argument I re-estimated a representative consumer model 100 times, each time with the observations randomly reshuffled along the time dimension. The results are reported in Appendix A, Table 11. Estimates of cable and satellite switching costs based on the reordered samples are not statistically significant, whereas price and quality coefficients appear to be similar to the estimates from the static version of the model.

[34] To accommodate larger state space I would need to reduce the number of points in each of the variables or the number of simulated consumers. This would increase the approximation error in the solution.

Appendix A.1 Estimates using different sets of IVs

Table 9: Robustness checks for different sets of IVs

variable	original	OLS-type	minimal IV
switching cost cab	1.177	2.205	0.997
	(0.208)	(0.161)	(0.340)
switching cost sat	1.470	2.760	1.395
	(0.210)	(1.018)	(0.256)
constant cab	0.506	-0.124	0.486
	(0.055)	(0.117)	(0.251)
constant sat	-0.993	-1.274	-1.160
	(0.152)	(0.199)	(0.639)
price	-0.073	-0.019	-0.076
	(0.002)	(0.001)	(0.012)
quality	0.273	0.132	0.285
	(0.009)	(0.004)	(0.013)
cab time dummies	yes	yes	yes
sat time dummies	yes	yes	yes
heterogeneity moments	yes	yes	yes
cab switching costs, USD'97	193.48	1370.18	158.44
sat switching costs, USD'97	241.64	1714.57	221.68
Hansen J-stat	96.061	481.461	64.73
p-val	0.000	0.000	0.002

Standard errors in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix A.2 Sensitivity with respect to initial conditions assumptions

To assess sensitivity of the estimates to the initial conditions, I conducted several robustness checks using the model with random coefficient on price (original estimation results are reproduced in column 1 of Table 10). First, I uniformly shifted the date of satellite entry from 1994 to 1995 and re-estimated the model with random coefficient on price. The results are reported in column 2 in Table 10. Second, I chose a satellite entry date for each market at random between 1994 and 1996, inclusive, and repeated estimation of the same model (columns 3). Finally, the last column 4 reports results for using a wider range for randomly assigned satellite entry, i.e., 1994-1997 inclusive.

Table 10: Sensitivity with respect to alternative initial conditions.

variable	(1)	(2)	(3)	(4)
		date of DBS entry		
	original	uniform 1995	rand. 1994-96	rand 1994-1997
switching cost cab	1.220	1.199	1.200	1.202
	(0.431)	(0.442)	(0.442)	(0.437)
switching cost sat	1.506	1.479	1.480	1.519
	(0.318)	(0.308)	(0.308)	(0.325)
constant cab	0.611	0.650	0.650	0.646
	(0.178)	(0.179)	(0.179)	(0.179)
constant sat	-0.986	-0.899	-0.899	-0.905
	(0.393)	(0.400)	(0.400)	(0.400)
price	-0.074	-0.076	-0.076	-0.076
	(0.009)	(0.009)	(0.009)	(0.009)
std(price)	0.012	0.013	0.013	0.013
	(0.004)	(0.004)	(0.004)	0.004)
quality	0.273	0.277	0.276	0.277
	(0.020)	(0.021)	(0.021)	(0.021)
cab & sat time-effects	yes	yes	yes	yes
heterogen. moments	no	no	no	no
cab switch. costs, $'97	197.84	188.35	188.51	189.03
sat switch. costs, $'97	244.22	232.40	232.48	239.01

Estimation results do not appear to be very sensitive to changes in the assumption about the date of entry by the DBS competitors. Of course, satellite entry is likely to be endogenous. For example, it is conceivable that DBS entered first into the markets with low cable market share. Allowing for an endogenous entry of satellite as well as endogenizing the number of products for cable providers are very interesting questions, which are left for further research.

Appendix A.3 Identification of switching costs, time-wise reordered data

Table 11: Results of 100 estimation runs, reshuffled data, representative consumer model.

variable	original		reordered			
	coef	(s.e.)	mean	sd	min	max
switching costs cab	1.031	(0.440)	-0.095	1.030	-1.998	1.322
switching costs sat	1.384	(0.280)	0.094	1.083	-1.442	1.982
price	-0.078	(0.010)	-0.123	0.009	-0.147	-0.113
quality	0.277	(0.025)	0.389	0.027	0.358	0.461

Appendix A.4 Consumer beliefs, robustness checks

This appendix presents results for alternative specifications of consumer beliefs. Flow utility values are obtained at the estimated parameter values from a representative consumer model (estimated using original specifications for consumer beliefs in equations (7) and (8) in the article). Original specifications are reproduced by columns (1) and (4). Columns (2) and (3) for cable and columns (5), (6), (7) are augmented with additional regressors to check if the model fit improves dramatically, in particular

specification (2) includes lagged price and quality levels for each product tier;

specification (3) includes both cable and satellite price and quality levels;
specification (5) includes average (across all cable firms) flow utility from cable providers;
specification (6) includes lagged price and quality levels for satellite;
specification (7) includes both cable and satellite price and quality levels.

Table 12: Robustness check of reduced-form consumer beliefs, more explanatory variables

variable	cable utility, δ_{ct}			satellite utility, δ_{st}			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
δ_{ct-1}	0.792***	0.766***	0.746***				
	(0.012)	(0.012)	(0.013)				
δ_{st-1}	0.059***	0.051***	0.094***	0.845***	0.861***	0.884***	0.880***
	(0.012)	(0.013)	(0.017)	(0.007)	(0.008)	(0.009)	(0.009)
p_{c0t-1}		-0.010***	-0.010***				-0.002**
		(0.002)	(0.002)				(0.001)
q_{c0t-1}		0.028***	0.040***				0.000
		(0.005)	(0.006)				(0.003)
p_{c1t-1}		-0.002*	-0.001				-0.005***
		(0.001)	(0.001)				(0.001)
q_{c1t-1}		0.021**	0.029***				0.027***
		(0.009)	(0.009)				(0.005)
p_{c2t-1}		0.000	-0.000				0.001
		(0.001)	(0.001)				(0.001)
q_{c2t-1}		-0.001	0.009				-0.010
		(0.011)	(0.011)				(0.007)
p_{s0t-1}			-0.232			-0.134*	-0.181**
			(0.142)			(0.081)	(0.085)
p_{s1t-1}			0.140			0.097*	0.132**
			(0.096)			(0.055)	(0.058)
p_{s2t-1}			0.043			0.021	0.029*
			(0.026)			(0.015)	(0.016)
q_{s0t-1}			-0.111			-0.021	-0.061
			(0.080)			(0.045)	(0.048)
q_{s1t-1}			-0.086			-0.119***	-0.131***
			(0.056)			(0.032)	(0.033)
q_{s2t-1}			0.077			0.048*	0.066**
			(0.047)			(0.027)	(0.028)
$\bar{\delta}_{ct-1}$					-0.170***		
					(0.035)		
const.	0.176***	0.311***	-0.176	0.189***	0.286***	-0.134	-0.306
	(0.010)	(0.032)	(0.491)	(0.005)	(0.020)	(0.281)	(0.294)
R^2	0.662	0.672	0.679	0.842	0.843	0.845	0.848

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 13: Robustness check of reduced-form consumer beliefs, extra lags

variable	δ_{ct} original		δ_{ct} with lags		δ_{st} original		δ_{st} with lags	
	coef	se	coef	se	coef	se	coef	se
δ_{ct-1}	0.792***	(0.012)	0.663***	(0.016)				
δ_{st-1}	0.059***	(0.012)	-0.040	(0.033)	0.845***	(0.007)	0.626***	(0.019)
δ_{ct-2}			0.206***	(0.016)				
δ_{st-2}			0.083**	(0.034)			0.236***	(0.019)
Constant	0.176***	(0.010)	0.158***	(0.013)	0.189***	(0.005)	0.247***	(0.007)
R-squared	0.662		0.685		0.842		0.850	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Appendix B A simple model of cable service provider

Simulating counterfactual experiments requires knowledge of the supply side parameters. Below I show that under a set of assumptions it is possible to solve for a reduced-form variable that contains information about the cost function, and which rationalizes choices of cable providers observed in the data. Then I assume that the variable and its evolution over time is invariant to the changes in market structure (elimination of switching costs and/or absence of the DBS competitors). Finally, a new optimal policy is calculated by jointly solving modified consumer and producer optimization problems. Below I detail the simulation exercise.

Each period a cable firm collects revenue equal to the subscription fee times the number of subscribers.[35] Producing quality, q_{ct}, is costly. Let $C(q_{ct}, Z_{ct}, Ms_{ct})$ denote the total cost of providing quality level q_{ct} to Ms_{ct} subscribers in the market of size M. The cost function depends on a vector of cost shifters Z_{ct}, which may contain exogenous unobserved service characteristics ξ_{ct}. I make the following assumption.

Assumption 4: *Cost structure of cable providers and consumer utility function satisfy the following set of restrictions:*

(i) Demand is given by a representative consumer with a linear per-period flow utility function (net of switching costs)

$$u_{igt} = \delta_{gt} + \epsilon_{igt}$$

where $\delta_{gt} = \alpha_g + \alpha_p p_{gt} + \alpha_q q_{gt} + \xi_{gt}$ *and* ϵ_{igt} *satisfies Assumption 1*

(ii) Total cost function of providing quality q_{ct} *is such that*

$$C(q_{ct}, Z_{ct}, Ms_{ct}) = Ms_{ct}\, C(q_{ct}, Z_{ct}),$$

i.e., marginal cost of providing quality q_{ct} *is constant in the number of subscribers. Moreover,* $C(q_{ct}, Z_{ct})$ *is not homogeneous of degree 1 in quality.*

(iii) Evolution of the cost shifters, Z_{ct}, *and exogenous unobserved service characteristics,* ξ_{ct}, *is invariant to the changes in market structure and consumer switching costs.*

Assumption 4(i) is consistent with the initial model setup except that I do not allow for persistent consumer heterogeneity. Part (ii) seems plausible because the service providers typically pay a fixed cost per subscriber to the networks they broadcast in the local markets. Nonhomogeneity of degree 1 in quality level is required to guarantee interior solution to the first-order conditions for optimal quality choice. Part (iii) emphasizes partial equilibrium nature of the counterfactual analysis.

Consider a single-product cable system.[36] Per-period profit function is given by

$$\begin{aligned}&\tilde{\pi}(\delta_c(p_{ct}, q_{ct}, \xi_{ct}), \delta_{st}, Z_{ct}, s_{ct-1}, s_{st-1}) = \\ &Ms\big(\delta_c(p_{ct}, q_{ct}, \xi_{ct}), \delta_{st}, s_{ct-1}, s_{st-1}\big)\big(p_{ct} - C(q_{ct}, Z_{ct})\big).\end{aligned} \tag{15}$$

[35] An important limitation of the model is the assumption that profit maximization from selling paid-television service is an independent problem of profit maximization from selling advertising. It is conceivable that the service providers solve a joint profit maximization problem accounting for the revenues from advertising. However, it is hard to build this into the model without making strong parametric assumptions and without additional data that can be informative about the parameters of interest.

[36] This is not an additional restriction but an outcome of Assumption 4, which implies that a profit maximizing service provider would rationally choose to offer only one product.

In the beginning of each period, the producer observes the current value of the "unobserved" (by us) product characteristics, ξ_{ct}, current period flow utility from satellite products, δ_{st}, and the realizations of the exogenous cost shifters, Z_{ct}. Hence, in the beginning of the period the producer has complete information about the current period profit function for any feasible choices of (p_{ct}, q_{ct}). Note that from the consumer standpoint holding δ_{ct} fixed any combinations of the (p_{ct}, q_{ct}) pair would generate the same utility level and beliefs about the future values of the flow utilities. From Assumption 4(i), $\delta_c(p_{ct}, q_{ct}, \xi_{ct})$ is strictly decreasing function of p_{ct} and, hence, there exists a well-defined inverse function

$$p = \delta^{-1}(\delta_{ct}, q_{ct}, \xi_{ct}).$$

Therefore, the producer choice variables can be redefined as (δ_{ct}, q_{ct}) instead of (p_{ct}, q_{ct}). Also, note that $\dfrac{\partial \delta^{-1}(\delta_{ct}, q_{ct}, \xi_{ct})}{\partial q_{ct}}$ is constant. Assuming that the producer maximizes expected present discounted value of future cash flows over an infinite horizon, the objective function at time τ is

$$\begin{gathered}\max_{\{\delta_{ct}, q_{ct}\}_{t=\tau}^{\infty}} \sum_{t=\tau}^{\infty} \beta^{t-\tau} \mathrm{E}\left[\pi(\delta_{ct}, q_{ct}, \xi_{ct}, \delta_{st}, Z_{ct}, s_{ct-1}, s_{st-1}) | \Omega_t^s\right] \\ (\xi_{ct}, \delta_{st}, Z_{ct}) \sim \mathrm{F}(\cdot|\Omega_t^s) \\ s_{ct} = s_c(\delta_{ct}, \delta_{st}, s_{ct-1}, s_{st-1}) \\ s_{st} = \delta_s(\delta_{ct}, \delta_{st}, s_{ct-1}, s_{st-1}),\end{gathered} \tag{16}$$

where expectation is with respect to future realizations of $(\xi_{ct}, \delta_{st}, Z_{ct})$, $s_c(\cdot)$ and $s_s(\cdot)$ are defined by the solution to consumer maximization problem, and Ω_t^s are factors that are informative about evolution of the exogenous variables. Dynamic optimization problem (16) can be written as a Bellman equation,

$$W(s_{ct-1}, s_{st-1}, \delta_{st}, \xi_{ct}, Z_{ct}, \Omega_t^s) =$$

$$\max_{\delta_{ct}, q_{ct}} \left\{ \begin{array}{c} s_c(s_{ct-1}, s_{st-1}, \delta_{st}, \delta_{ct})\big(\delta^{-1}(\delta_{ct}, q_{ct}, \xi_{ct}) - C(q_{ct}, Z_{ct})\big) + \\ \beta\, \mathrm{E}\big[W(s_{ct}, s_{st}, \delta_{st+1}, \xi_{ct+1}, Z_{ct+1}) | \Omega_t^s\big] \end{array} \right\}. \tag{17}$$

Recall that conditional on the policy choice, δ_{ct}, the choice of quality, q_{ct}, does not have any dynamic implications. From the first-order conditions for optimal quality choice,

$$\mathrm{FOC}[q_{ct}] : \frac{\partial \delta^{-1}(\delta_{ct}, q_{ct}, \xi_{ct})}{\partial q_{ct}} - \frac{\partial C(q_{ct}, Z_{ct})}{\partial q_{ct}} = 0, \tag{18}$$

it is clear that $q_{ct}^* = q(Z_{ct})$ does not depend on the optimal choice of a dynamic control but is defined by a vector of cost-shifters Z_{ct} (which may include ξ_{ct}). Therefore, the producer dynamic programming problem can be redefined by partially maximizing with respect to q_{ct} (aka "static-dynamic" breakdown),

$$W(s_{ct-1}, s_{st-1}, \delta_{st}, \xi_{ct}, Z_{ct}, \Omega_t^s) =$$

$$\max_{\delta_{ct}} \left\{ \begin{array}{c} s_c(s_{ct-1}, s_{st-1}, \delta_{st}, \delta_{ct})\Big(\dfrac{1}{\alpha_p}\big(\delta_{ct} - \alpha_c - \alpha_q q^*(Z_{ct}) - \xi_{ct}\big) - \tilde{C}(Z_{ct})\Big) \\ +\beta\, \mathrm{E}\big[W(s_{ct}, s_{st}, \delta_{st+1}, \xi_{ct+1}, Z_{ct+1}, \Omega_{t+1}^s) | \Omega_t^s\big] \end{array} \right\}.$$

Let

$$\Psi(\xi_{ct}, Z_{ct}) \equiv \frac{1}{\alpha_p}\big(- \alpha_c - \alpha_q q^*(Z_{ct}) - \xi_{ct}\big) - \tilde{C}(Z_{ct}) \tag{19}$$

Then

$$W(s_{ct-1}, s_{st-1}, \delta_{st}, \xi_{ct}, Z_{ct}, \Omega_t^s) =$$

$$\max_{\delta_{ct}} \left\{ \begin{array}{c} s_c(s_{ct-1}, s_{st-1}, \delta_{st}, \delta_{ct})\big(\frac{1}{\alpha_p}\delta_{ct} + \Psi(\xi_{ct}, Z_{ct})\big) \\ +\beta\, \mathrm{E}\big[W(s_{ct}, s_{st}, \delta_{st+1}, \xi_{ct+1}, Z_{ct+1}, \Omega_{t+1}^s)|\Omega_t^s\big] \end{array} \right\}.$$

In order to reduce the dimensionality of the state space, I make another major simplifying assumption.

Assumption 5: *The producer perceives the current period value of function $\Psi(\xi_{ct}, Z_{ct})$ as a sufficient statistic for the distribution over its future values,*

$$\mathrm{P}(\Psi_{ct+1}|\Omega_t^s, \Psi_{ct}) = \mathrm{P}(\Psi_{ct+1}|\Psi_{ct})$$

with the empirical version given by

$$\Psi_{ct+1} = \beta_0 + \beta_1\Psi_{ct} + \sigma_H\, v_{ct+1}$$

Obviously, the assumption is quite strong as it imposes a first-order Markov process restriction on the random variable obtained as a transformation of a vector of random variables with unknown distribution.[37] However, this assumption reduces the producer state space considerably, which makes it feasible to numerically solve the dynamic programming problem. Under Assumption 5, $\Omega_t^s = (\delta_{st}, \Psi_{ct})$, where the evolution of utility from satellite satisfies equation (7). In practice, I solve the following modified producer Bellman equation,

$$W(s_{ct-1}, s_{st-1}, \delta_{st}, \Psi_{ct}) = \max_{\delta_{ct}} \left\{ \begin{array}{c} s_c(s_{ct-1}, s_{st-1}, \delta_{st}, \delta_{ct})\big(\frac{1}{\alpha_p}\delta_{ct} + \Psi_{ct}\big) \\ +\beta\, \mathrm{E}\big[W(s_{ct}, s_{st}, \delta_{st+1}, \Psi_{ct+1}|\delta_{st}, \Psi_{ct}\big] \end{array} \right\}.$$

Note that by Assumption 4(iii), realizations of $\Psi_{ct} = \Psi(\xi_{ct}, Z_{ct})$ stay constant across various scenarios. Therefore, I can solve for Ψ_{ct} at a finite number of points in the producer state space to recover a sequence of numbers, $\{\Psi_{ct}\}_{t=0,\ldots,T}$ that rationalizes the sequence of policy choices, $\{\delta_{ct}\}_{t=0,\ldots,T}$, observed in the data. This can be done by first solving for the optimal producer policy $\delta_{ct} = \delta(s_{ct-1}, s_{st-1}, \delta_{st}, \Psi_{ct})$ and then "inverting-out" Ψ_{ct} values by matching model predictions to the actual data.[38]

The solution algorithm is iterative and consists of several steps. An empirical version of the transition of Ψ_{ct} is initialized by choosing parameters $(\beta_0, \beta_1, \sigma_H)$. Transition parameters for the state variable δ_{st} are set at the values obtained from the demand side estimation. I begin by computing optimal δ_{ct} policy at a finite number of grid points in the state space. Then the predicted policy is "inverted" to obtain a set of Ψ_{ct} numbers rationalizing choices

[37]Note that optimal policy δ_{ct} is likely to violate the first-order Markov assumption. In the model, the "information exchange" between consumers and producers is restricted to the AR(1) processes. Therefore, consumers are modeled as boundedly rational, i.e., rationality is restricted to what can be captured by the simple reduced-form specifications of beliefs. Unfortunately, without estimating a joint supply-demand model it is hard to assess the role of the bounded rationality assumption.

[38]For a simple transition process, like the one assumed for Ψ_{ct}, it can be shown that the optimal policy is strictly monotonic in Ψ_{ct}.

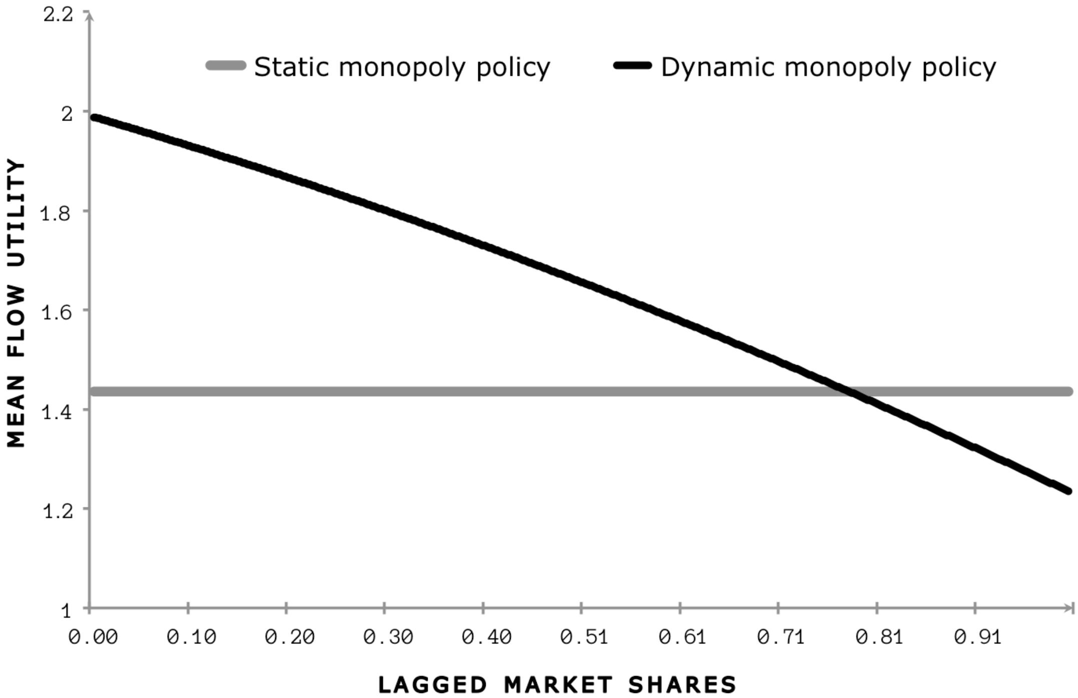
actually observed in the data. Then transition parameters for Ψ_{ct} are updated. The iterations are repeated until complete convergence on both the producer value function and the resulting optimal policy. The recovered Ψ_{ct} values in each market provide enough information to simulate several counterfactual scenarios.

For the dynamic monopoly case, when there is no DBS but there are consumer switching costs, I redefine the consumer and producer dynamic programming problems and solve them iteratively by updating consumer beliefs in equation (6) after solving for the optimal producer policy. A new equilibrium sequence of δ_{ct} in each market is obtained when complete convergence on (1) consumer Bellman, (2) producer Bellman, and (3) optimal producer policy is reached.

It is worth noting that Ψ_{ct} encapsulates information not only on the costs of providing quality but also on the resulting choices of quality itself as is apparent from equation (19). Therefore, it would be impossible to apportion differences in the counterfactual cable policy, defined as δ_{ct}, separately into price and quality levels. In this setup, the evolution of optimal quality is exogenous and not separately identifiable of the evolution in the cost function, however, switching costs and competition do effect optimal choices of cable firms by altering their price choices. An alternative would be to assume parametric form of the cost function and to deal with a much larger object (cost shifters) in the state space.

Figure 1 illustrates the differences in the optimal policy, δ_{ct}, for dynamic and static monopolists as functions of their last period market shares.

Figure 1: Optimal δ_{ct} for a dynamic and static monopolist, holding Ψ_{ct} fixed.



As expected, the optimal dynamic monopolist's policy can be larger or smaller than the optimal policy of a static monopolist depending on the market share. At the early stages, the investment motive dominates and prices are lower (utilities are higher) allowing to "buildup" the consumer base. With sufficiently large market share a monopolist will raise prices, harvesting the rents.

References

Ackerberg, D. and Rysman, M. (2005). Unobserved product differentiation in discrete-choice models: estimating price elasticities and welfare effects. *The RAND Journal of Economics*, 36(4):771–788.

Aguirregabiria, V. and Nevo, A. (2013). Recent developments in empirical IO: dynamic demand and dynamic games. In Acemoglu, D., Arellano, M., and Dekel, E., editors, *Advances in Economics and Econometrics*, volume 3, econometric society monographs 2, pages 53–122. Cambridge University Press.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics*, 25:242–262.

Berry, S. and Haile, P. (2013). Identification in differentiated products markets using market level data. Unpublished manuscript. Yale University.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63:841–890.

Berry, S. and Pakes, A. (2007). The pure characteristics demand model. *International Economic Review*, 48(4):1193–1225.

Chamberlain, G. (1992). Comment: sequential moment restrictions in panel data. *Journal of Business and Economic Statistics*, 10(1):20–26.

Chu, C. S. (2007). The effect of satellite entry on product quality for cable. Unpublished dissertation, Stanford University.

Crawford, G. S. (2000). The impact of the 1992 cable act on household demand and welfare. *RAND Journal of Economics*, 31:442–449.

Crawford, G. S. (2005). The discriminatory incentives to bundle in the cable television industry. Mimeo, University of Arizona.

Crawford, G. S. and Shum, M. (2007). Monopoly quality degradation and regulation in cable television. *Journal of Law and Economics*, 50:181–219.

Crawford, G. S. and Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review*, 102(2):643685.

Dube, J.-P., Hitsch, G., and Rossi, P. (2009). Do switching costs make markets less competitive? *Journal of Marketing Research*, 46(4):435–445.

Dube, J.-P., Hitsch, G., and Rossi, P. (2010). State dependence and alternative explanations for consumer inertia. *RAND Journal of Economics*, 41(3):417–445.

Farrell, J. and Klemperer, P. (2007). Coordination and lock-in: competition with switching costs and network effects. In Armstrong, M. and Porter, R., editors, *Handbook of Industrial Organization*, volume 3, chapter 31, pages 1967–2072. Elsevier B.V.

FCC (1998). Annual assessment of the status of competition in markets for the delivery of video programming. Fourth Annual Report 97-423, Federal Communications Commission.

Fox, J., Kim, K., Ryan, S., and Bajari, P. (2012). The random coefficients logit model is identified. *Journal of Econometrics*, 166:204–212.

Goolsbee, A. and Petrin, A. (2004). The consumer gains from Direct Broadcast Satellites and the competition with cable TV. *Econometrica*, 72:351–381.

Gowrisankaran, G. and Rysman, M. (2012). Dynamics of consumer demand for new durable goods. *Journal of Political Economy*, 120(6):1173–1219.

Handel, B. (2013). Adverse selection and inertia in health insurance markets: when nudging hurts. *American Economic Review*, 103(7):2643–2682.

Hendel, I. and Nevo, A. (2006). Measuring the implications of sales and consumer inventory behavior. *Econometrica*, 74:1637–1673.

Kim, J. (2006). Consumers' dynamic switching decisions in the cellular service industry. Working Paper 06-24, The Networks, Electronic Commerce, and Telecommunications Institute.

Klemperer, P. (1987a). The competitiveness of markets with switching costs. *RAND Journal of Economics*, 18:138–150.

Klemperer, P. (1987b). Markets with consumer switching costs. *The Quarterly Journal of Economics*, 102:375–394.

Magnac, T. and Thesmar, D. (2002). Identifying discrete decision processes. *Econometrica*, 70(2):801–816.

Manski, C. F. (1993). Dynamic choice in social settings. Learning from the experiences of others. *Journal of Econometrics*, 58:121–136.

Melnikov, O. (2013). Demand for differentiated durable products: the case of the U.S. computer printer market. *Economic Inquiry*, 51(2):1277–1298.

Nosal, K. (2012). Estimating switching costs for Medicare Advantage plans. Unpublished manuscript. University of Arizona.

Rust, J. (1987). Optimal replacement of GMC bus engines: an empirical model of Harold Zurcher. *Econometrica*, 55(5):999–1033.

Rust, J. (1994). Structural estimation of Markov decision processes. In Engle, R. and McFadden, D., editors, *Handbook of Econometrics*, volume 4, chapter 51, pages 3081–3143. Elsevier Science B.V.

Schiraldi, P. (2011). Automobile replacement: a dynamic structural approach. *RAND Journal of Economics*, 42(2):266–291. Summer.

Shcherbakov, O. (2009). The effect of consumer switching costs on market power of cable television providers. Unpublished manuscript.

Sudhir, K. and Yang, N. (2014). Exploiting the choice-consumption mismatch: a new approach to disentangle state dependence and heterogeneity. Cowles Foundation Discussion Paper No. 1941.

Yang, Y. H. (2010). Identification of consumer switching behavior with market level data. Unpublished manuscript. UCLA.