The effect of consumer switching costs on market power of cable television providers

Oleksandr Shcherbakov

November 2009

[Preliminary and incomplete]

Abstract

This paper empirically evaluates the effect of consumer switching costs on market power of cable television providers. To do so I propose an algorithm to estimate supply side parameters when the demand side is represented by consumers with persistent heterogeneity in tastes and state-dependent utility. Under such conditions, regardless of whether consumers are forward-looking or behave in a myopic way, the demand schedule introduces dynamics into the producer problem. Presence of multiple state variables makes full solution to the producer dynamic programming problem difficult or computationally infeasible. Therefore, I estimate supply side parameters from the optimality conditions for the dynamic controls. Using data from the paid television markets in US, I estimate cost functions of the cable television providers. In these markets, persistent consumer heterogeneity and state dependence due to switching costs requires producers to keep track of the entire distribution of consumer types which enters states space of a dynamic programming problem. Preliminary estimates of the consumer switching costs are $149 and $238 for cable and satellite providers respectively. The estimates of the supply side parameters imply average cost of providing service per subscriber of $2.19 and average per subscriber price-cost margin of $16.52. Counterfactual simulations suggest that without switching costs cable prices would be 28 percent higher with satellite competitors and 51 percent higher in case of static monopoly scenario.

1 Introduction

This paper pursues two objectives. The first one is to assess the economic effects of consumer switching costs in the paid-television industry on market structure and social welfare. To accomplish this task I need to evaluate the impact of switching costs on the optimal choices of prices and quality levels by cable television providers. This, in turn, requires knowledge of the costs structure of the cable firms.

Estimation of the supply side parameters becomes a challenging task if consumer utility exhibits state-dependence generated by the non-trivial switching costs. State-dependent demand, in turn, requires considering price and quality choices by the producers in a dynamic perspective. This is true regardless of whether consumers make myopic or forward-looking decisions. With multiple consumer types and/or many products in the market the state space of the producer dynamic programming problem becomes very large. Large state space renders traditional solutions to the producer problem computationally infeasible. Hence, the second objective of this paper is to develop a methodology that makes estimation of the supply side parameters possible.

In order to overcome the large state space problem I suggest estimation from the first-order conditions for dynamic controls similar (but not identical) to Berry and Pakes (2001). A modified generalized instrumental variables technique for non-linear rational expectations models originally proposed by Hansen and Singleton (1982) could be used to estimate parameters of the cable company cost function. Presence of multiple state variables that are simultaneously shifted by a single dynamic control makes it impossible to express the first-order conditions in terms of primitives of the model, i.e. derivatives of the per-period reward and transition functions. To obtain partial derivatives of the value function with respect to each state I use forward simulation approach developed in Hotz et al. (1994) for single-agent decision problems and extended in Bajari et al. (2007) to a dynamic games environment.

Simulation approach consists of three steps. In the first step, I estimate demand side model and recover the distribution of the consumer types across market shares and unobserved cable and satellite service characteristics. In the second step, I estimate policy functions defined on the producer state space and the law of motion for the exogenous state variables. In the third step, parameters of the cost function are estimated from the first-order conditions for the producer dynamic programming problem. The main difference from Bajari et al. (2007) is in the first step, which is necessary to recover unobserved producer state variables.

Similar to Bajari et al. (2007), there are considerable computation benefits if per-period producer reward function is linear in parameters. In this case, computationally intense simulation of the derivatives of the value function needs to be conducted only once for a set of basis functions. When estimating supply-side parameters the vector of basis functions is simply scaled by the current values of the parameter vector. This makes estimation procedure very fast and comparable in terms of computation time to a standard GMM procedure for non-linear models.

To estimate parameters of the demand and supply models I use data on 564 U.S. cable systems in 1992-2002. Preliminary estimates of the demand side parameters suggest switching costs of \$149 and \$238 for cable and satellite providers respectively. At this stage, the demand side estimates cannot reject a representative consumer model, although mean levels of switching costs as well as price and quality coefficients are estimated precisely. Supply-side estimates imply the average cost of providing service per subscriber of about $2.19 and average price-cost margin of $16.52. Under counterfactual scenario where there are no switching costs and satellite policy as well as the values of own cable companies' cost shifters remain unchanged cable prices are estimated to be 28 percent higher than the observed level. In a static monopoly scenario, where there are no satellite competitor and no switching costs cable prices are on average 51 percent higher than the observed ones.

The remainder of the paper is organized as follows. Section 2 discusses related literature. Sections 3 and 4 provide institutional details for the paid-television industry in U.S. and describe data used for estimation. In sections 5 and 6 I describe demand and supply models and corresponding estimation strategies and outline differences in estimation algorithm for myopic and dynamic consumers. Section 7 discusses instrumental variables and identification issues and section 8 presents preliminary estimation results. Section 9 outlines strategy for simulating counterfactual scenarios of the industry evolution and section 10 concludes.

2 Related literature

Despite a large theoretical literature on switching costs, there are only a few empirical studies. Several examples of industries with consumer switching costs include banking (Sharpe (1997), Kiser (2002), Kim et al. (2003)), auto insurance (Israel (2003)), airline (in relation to frequent flyer programs; Borenstein (1992)), long-distance telephone service (Knittel (1997)), and retail electricity industries (Salies (2005), Sturluson (2002)).

The limited number of empirical studies on the topic may reflect the difficulty of measuring switching costs, which are not directly observable in the data. One of the widely cited papers in the field is Shy (2002). He suggests a framework for quick and easy estimation of switching costs. Under a set of assumptions, the author shows how switching costs can be directly inferred from observed prices and market shares. Unfortunately, the underlying assumptions are very strong, and include homogeneous products and static behavior on both the demand and supply side. Two other papers that attempt to empirically measure switching costs in the framework of dynamically optimizing consumers are Schiraldi (2006) and Ho (2009). Both papers use modified versions of the nested fixed point estimation algorithm developed in Gowrisankaran and Rysman (2007) for estimation of dynamic demand for durable goods.

While I specify and estimate a model of consumer behavior in markets with switching costs this is not the primary objective of the present paper. The major focus is on a methodology that allows estimating supply side parameters, namely parameters of the cost function, in presence of state dependent consumer utility. State dependence on the demand side calls for a forward looking behavior on the supply side. In presence of many state variables traditional solution to the producer dynamic programming problem becomes computationally infeasible. Hence, alternative estimation approaches are needed.

There are only a few empirical papers that attempt to address the question. The first one is Dube et al. (2008) where the authors estimate the demand side with many heterogeneous consumers. To solve for the dynamically optimal pricing policy they assume a stationary long-run equilibrium, which allows using Euler equation framework. It is worth noting that this is rather a computational theory paper (in its supply side part) that calculates alternative optimal pricing policies under a set of very restrictive assumptions, including steady state and no uncertainty on the producer side. Another paper in the area is Che et al. (2006) who deal with the problem of large state space by imposing bounded rationality assumption on the firms behavior, when the distribution of the consumer types across products is summarized by its first moments.

It is worth noting that one of the possible solutions to the problem of many state variables in macro literature was suggested in Krusell and Smith (1998). The idea is that the distribution of consumer types could be approximated reasonably well with only a finite set of its first moments. However, in case of several products state-dependence requires keeping track of these moments for each product. Hence, with the increase in the number of products solution to the producer dynamic programming problem quickly becomes infeasible.

3 Institutional details

Cable television, formerly known as Community Antenna Television or CATV, emerged in the late 1940s in Arkansas, Oregon and Pennsylvania to deliver broadcast signals to the remote areas with poor over-the-air reception.[1] In these areas homes were connected to the antenna towers located at the high points via cable network. Starting with 70 cable systems serving about 14,000 subscribers in 1952, a decade later almost 800 cable systems served about 850,000 subscribers (ibid.). According to FCC (2000), by October 1998 the number of cable systems reached 10,700 providing service to more than 65 million subscribers in 32,000 communities.

The ability of cable systems to "import" distant signals imposed a significant competitive pressure on the local television stations, which eventually has led to a regulatory restrictions by the FCC on the programming content of cable companies first introduced in 1965-66. Gradual deregulation of cable industry began in the early 1970s, which when accompanied with the development of satellite communication technology has led to an emergence of national networks (e.g. HBO in 1972, Showtime in 1976, ESPN in 1979), which programming was distributed by satellite to cable systems nationwide. Considerable investment activity in the industry resumed after the 1984 Cable Communications Policy Act that established a more favorable regulatory environment. In 1992, continuing increase in the cable prices resulted in another waive of regulatory intervention, when Congress enacted the Cable Televi-

[1] See National Cable & Telecommunications Association (NCTA), `http://www.ncta.com/About/About/HistoryofCableTelevision.aspx` (accessed March 01, 2009).

sion Consumer Protection and Competition Act of 1992. Despite new regulatory restrictions the industry continued its growth. At about the same time, a new competing technology - direct broadcast satellite - challenged previously "exclusive" cable programming.

The 1996 Telecommunications Act introduced a dramatic change in the public policy for telecommunication services towards further deregulation of the industry. A major upgrade of the cable distribution networks amounting to about $65 billion of investment between 1996 and 2002 resulted in an increased amount of hybrid networks of fiber optic and coaxial cable (see an overview of cable history by NCTA). These high capacity networks allowed for a multi-channel video, two-way voice, high-speed internet access and high definition digital video services. In the late 90s, most cable systems had capacity ranging between 36 and 60 channels and some offered more than 100 channels. Most cable subscribers receive service from a system offering more than 54 channels (FCC (2000)). In some cases, cable companies created own programming as well as provided leased access channels to those wishing to show specific programs. Other home services that are possible using the two-way transmission cable networks include video on demand, interactive TV, electronic banking, shopping, utility meter reading, etc.

Until the 1990s, local cable systems were effectively natural monopolies as they faced virtually no competition except in a few cases of "overbuilt" systems where the same location was served by more than one cable company. Competition from the C-Band satellite (a predecessor to today's DBS systems) was very limited because of extremely high setup costs. DBS service was launched in the early 90s and originally was popular mostly in rural areas where cable service did not exist. Since then the number of subscribers of DBS providers has experienced rapid growth.

Table 1: DBS penetration rates in 2001-2004

	2001	2004	Change
Rural	26%	29%	12%
Suburban	14%	18%	29%
Urban	9%	13%	44%

Source: GAO report to the U.S. Senate, April 2005

Due to technological restrictions, DBS cannot match most of the supplementary services offered by the cable companies. Until recently, some differences in the programming content were induced by industry regulation. Prior to 1999, when the Congress enacted the Satellite Home Viewer Improvement Act, DBS carriers were not allowed to broadcast local channels. In many cases this was considered a competitive disadvantage of satellite providers (see FCC 4th annual report on competition in markets for video programming, as of January 13, 1998).

DBS and cable operators use different quality and price setting strategies. While each cable system makes pricing and quality decisions locally, satellite operators set these variables at the national level. It is conceivable though that there are other factors, like customer service as well as landscape and weather conditions that effect the quality of reception and attractiveness of the satellite service at regional level.

Switching costs in the television industry are primarily transactional. They include not only upfront equipment and installation fees but also hassle costs. Even though many cable companies do not require purchasing equipment (it is rather rented by the cable subscribers), anecdotal evidence suggests that the cost of returning rented equipment may be substantial. In what follows I use terms “start-up costs” and “switching costs” interchangeably. Another component of the switching costs are “shopping costs” associated with the purchase of supplementary services like telephone and internet.

4 Data

In order to measure the economic effects of switching costs I use data from the US paid television industry in 1992-2002. Most variables for the cable providers come from the Warren’s Factbook editions. Satellite data was collected from the internet sources and covers 1997-2002.

For the empirical estimation part of the paper I define market to be an area franchised to the cable company. In all of the markets used for estimation there is only a single cable system, i.e. none of the “overbilders” are included.

Even though cable and satellite providers often offer more then one programming tier, I maintain a single service assumption.[2] Thus, for cable it is either Basic or Expanded Basic depending on the number of subscribers and for satellite it is Total Choice (DIRECTV) tier.

For estimation I used a sample of 564 cable systems that have reliable data on market shares, price and quality variables for 1992-2002. Below I describe several issues with the data.

Satellite penetration rates are available only at the Designated Market Area (DMA) level. In order to compute satellite market share for each of the more narrowly defined markets I used an assumption similar to one in Chu (2007). Within a DMA satellite subscribers constitute a constant proportion of the non-cable subscribers. Define

$$R_{kt} = \frac{\#satsubs_{kt}}{M_{kt} - \#cabsubs_{kt}}$$

where k and t are DMA and time subscripts; $\#satsubs_{kt}$ is the number of satellite subscribers; M_{kt} is the total number of households; and $\#cabsubs_{kt}$ is the number of cable subscribers. Then satellite market share in market j located in DMA k is computed as

$$s^s_{jt} = (1 - s^c_{jt})R_{kt}$$

[2]There are no conceptual difficulties in including multiple tiers. It is rather data limitations, when the number of missing observations for more advanced tiers is too large.

The rationale for this assumption is related to the timing of the entry by DBS. In the first place, satellite providers target areas where there is no alternative cable paid-television service or where the cable share is small. Therefore, one can expect that within the same DMA satellite penetration is relatively larger in the areas franchised to the cable systems with smaller market shares. Typically, satellite penetration is greater in rural and suburban than in the densely populated urban areas where cable companies have greater market shares.[3] Geographic variation in satellite penetration rates is supported by official statistics (see Section 3).

Besides, DMA level satellite penetration rates are available only from 1997. In order to deal with the initial conditions problem (as discussed in section 6.1), I imputed satellite shares at the local markets level by using the national level dynamics. For this reason, to form the empirical moment conditions I use data in 1997-2002 only.[4]

Another important question is the definition of the quality of programming content offered by a particular provider. Using number of channels offered as a proxy for quality of the system's programming may be problematic. In particular, such a proxy would not capture changes in the programming composition, holding the number of channels constant. In many cases, the data reveal that a lot of variation in the quality variable is due to the change in the composition of channels rather than due to change in the number of channels. In order to control for different compositions of channels I used data on the average cost of each channel charged by the television networks. Channels with unknown or zero costs were assigned a cost of $0.01.

Price data for cable and satellite services was adjusted using consumer price index with 1997 as the base year. Hence, any monetary equivalents computed in this paper are in 1997 prices. Summary statistics for the key variables is listed in the table 2

5 Model

In this section I discuss demand and supply side models of the paid television industry. On the demand side, consumers are persistently heterogeneous with respect to their switching costs and tastes for service characteristics. Presence of switching costs results in state-dependence when each consumer's last period choice affects current utility. State-dependence, in turn, calls for the forward-looking consumer behavior. However, it is not clear if the gains

[3]Another reason to expect lower satellite penetration in urban areas is the necessity to locate receiver (dish) in a place that guarantees open access to the orbital satellite. In urban areas it was harder due to the presence of multistory buildings that may impede receiving satellite beams. Besides in multi-unit structures up until recently to install a dish a resident must obtain permission of the home owner, which was not always an easy task.

[4]Data from 1992-1996 are used only to approximate the initial distribution of consumer types across market shares.

Table 2: Descriptive statistics for the key variables, 564 systems, 6054 observations

Variable	min	max	mean	med	s.d.
Market share (cable)	0.011	0.867	0.426	0.438	9.157
Market share (satellite)	0.001	0.364	0.083	0.074	0.065
Price (cable)	3.551	34.20	19.20	20.52	5.869
Price (satellite)	24.80	34.74	29.85	29.99	3.687
Quality (cable)	0.02	9.570	3.338	3.06	1.558
Quality (satellite)	4.49	10.42	5.826	4.49	2.087
Capacity (cable)	10	134	39.78	36	13.58
Miles coaxial lines (cable),'000	0.001	4.491	0.111	0.020	0.301

Source: own calculations

from a dynamically optimal behavior justifies costs of finding such an optimal strategy.[5] In the present paper, I focus primarily on the supply side and maintain an assumption of consumer myopia for the reasons of computational simplicity. However, whenever relevant I provide quick asides on the modifications necessary to incorporate dynamic considerations into the demand-side model. Dynamic demand estimates using the same data can be found in Shcherbakov (2008). Moreover, as I discuss below, the suggested supply-side estimation strategy need not change under alternative assumptions on the consumer rationality as long as the estimation of the demand relationship does not depend on the parameters of the supply side.

5.1 Demand side: myopic consumers

Let $J = \{o, c, s\}$ denote consumer choice set consisting of cable, c, satellite, s, and an outside option of no paid TV, o. There is a number of persistently heterogeneous consumer types whose (time-invariant) preferences are given by iid draws from a distribution known up to a parameter vector. Within each type there is a continuum of non-persistently heterogeneous consumers. With some abuse of notation both types of heterogeneity will be subscribed with i.

Cable and satellite television is characterized by two observable characteristics: monthly subscription fee, p_{jt}, and quality index, q_{jt}. In addition, there is a scalar, ξ_{jt}, representing service characteristics observed by the market participants, but unobserved by an econometricians.

Let $a_{it} \in J$ denote consumer choice in period t. First time subscription to service j is associated with time-invariant monetary and hassle costs expressed in utility units, η_{ij}.

[5]Quantifying such gains goes beyond the scope of the present paper and remains a question for further research.

Assume that disconnection is costless, i.e. $\eta_{io} = 0\ \forall i$. Persistent consumer heterogeneity in tastes for service has several dimensions. In addition to the type-specific switching costs, each consumer type has idiosyncratic taste for service and different price sensitivity.

Assume that consumer per-period utility function is linear and given by

$$
\begin{aligned}
u_{ijt} &= \begin{cases} -\eta_{ij}+ & \alpha_{ij} + \alpha_i^p p_{jt} + \alpha^q q_{jt} + \xi_{jt} + \epsilon_{ijt}, \text{ if } a_{it} \neq a_{it-1}, \\ & \alpha_{ij} + \alpha_i^p p_{jt} + \alpha^q q_{jt} + \xi_{jt} + \epsilon_{ijt}, \text{ otherwise} \end{cases} \\
&= \begin{cases} -\eta_{ij}+ & \alpha_{ij} + \tilde{\alpha}_i p_{jt} + \delta_{jt} + \epsilon_{ijt}, \text{ if } a_{it} \neq a_{it-1}, \\ & \alpha_{ij} + \tilde{\alpha}_i p_{jt} + \delta_{jt} + \epsilon_{ijt}, \text{ otherwise} \end{cases} \\
&= \begin{cases} -\eta_{ij}+ & \delta_{ijt} + \epsilon_{ijt}, \text{ if } a_{it} \neq a_{it-1}, \\ & \delta_{ijt} + \epsilon_{ijt}, \text{ otherwise} \end{cases}
\end{aligned} \tag{1}
$$

where $\delta_{jt} = \bar{\alpha}^p p_{jt} + \alpha^q q_{jt} + \xi_{jt}$, $\delta_{ijt} = \alpha_{ij} + \tilde{\alpha}_i^p p_{jt} + \delta_{jt}$, $\bar{\alpha}^p$ is the population mean price sensitivity, $\tilde{\alpha}_i^p$ are each type deviations from the mean, and $\epsilon_{it} = (\epsilon_{iot}, \epsilon_{ict}, \epsilon_{ist}) \overset{iid}{\sim} F_\epsilon(\cdot)$ represent (non-persistent) heterogeneity within each discrete consumer type. I normalize utility from the outside option (over-the-air television) to zero, i.e. $u_{iot} = \epsilon_{iot}$.

Assume that consumers are myopic, i.e. their choices are based on the current period utility and do not account for the future evolution of service characteristics. Define the following probabilities for consumer type i

$$
\begin{aligned}
Pr_i(c \to c) &= \Pr(\delta_{ict} + \epsilon_{ict} \geq -\eta_{is} + \delta_{ist} + \epsilon_{ist}, \delta_{ict} + \epsilon_{ict} \geq \epsilon_{iot}) \\
Pr_i(s \to c) &= \Pr(-\eta_{ic} + \delta_{ict} + \epsilon_{ict} \geq \delta_{ist} + \epsilon_{ist}, -\eta_{ic} + \delta_{ict} + \epsilon_{ict} \geq \epsilon_{iot}) \\
Pr_i(o \to c) &= \Pr(-\eta_{ic} + \delta_{ict} + \epsilon_{ict} \geq -\eta_{is} + \delta_{ist} + \epsilon_{ist}, -\eta_{ic} + \delta_{ict} + \epsilon_{ict} \geq \epsilon_{iot}) \\
Pr_i(c \to s) &= \Pr(-\eta_{is} + \delta_{ist} + \epsilon_{ist} \geq \delta_{ict} + \epsilon_{ict}, -\eta_{is} + \delta_{ist} + \epsilon_{ist} \geq \epsilon_{iot}) \\
Pr_i(s \to s) &= \Pr(\delta_{ist} + \epsilon_{ist} \geq -\eta_{ic} + \delta_{ict} + \epsilon_{ict}, \delta_{ist} + \epsilon_{ist} \geq \epsilon_{iot}) \\
Pr_i(o \to s) &= \Pr(-\eta_{is} + \delta_{ist} + \epsilon_{ist} \geq -\eta_{ic} + \delta_{ict} + \epsilon_{ict}, -\eta_{is} + \delta_{ist} + \epsilon_{ist} \geq \epsilon_{iot})
\end{aligned}
$$

Let s_{ijt} denote a share of consumer type i choosing product j in period t. Then in any $t > 0$ current share of consumer type i subscribed to the cable service is given by

$$
s_{ict} = s_{ict-1} \cdot Pr_i(c \to c) + s_{ist-1} \cdot Pr_i(s \to c) + (1 - s_{ict-1} - s_{ist-1}) \cdot Pr_i(o \to c) \tag{2}
$$

Similarly, current share of consumer type i subscribed to the satellite service is given by

$$
s_{ist} = s_{ist-1} \cdot Pr_i(s \to s) + s_{ict-1} \cdot Pr_i(c \to s) + (1 - s_{ict-1} - s_{ist-1}) \cdot Pr_i(o \to s) \tag{3}
$$

Assumption 1: *Consumer heterogeneity parameters ϵ_{it} are represented by iid draws from a* *distribution known up to a parameter vector, i.e.*

$$\epsilon_{ijt} \overset{iid}{\sim} \textit{Extreme Value Type } 1, \textit{ with density}$$
$$f(\epsilon_{ijt}) = \exp(-\epsilon_{ijt}) \exp(-\exp(-\epsilon_{ijt}))$$

By assumption 1 the share of type i choosing cable in the current period is given by

$$\begin{aligned} s_{ict} =& s_{ict-1} \cdot \frac{\exp(\delta_{ict})}{1 + \exp(\delta_{ict}) + \exp(-\eta_{ist} + \delta_{ist})} \\ &+ s_{ist-1} \cdot \frac{\exp(-\eta_{ic} + \delta_{ict})}{1 + \exp(-\eta_{ic} + \delta_{ict}) + \exp(\delta_{ist})} \\ &+ (1 - s_{ict-1} - s_{ist-1}) \cdot \frac{\exp(-\eta_{ic} + \delta_{ict})}{1 + \exp(-\eta_{ic} + \delta_{ict}) + \exp(-\eta_{ist} + \delta_{ist})} \end{aligned}$$

Individual satellite shares for each type i can be calculated in a similar way.

Let me postpone derivation of the aggregate market shares and discuss an alternative model where consumers are forward-looking.[6]

5.2 Demand side: forward-looking consumers

In this section, I discuss a possibility to incorporate dynamic considerations into the consumer behavior. As long as the formulation of the consumer dynamic programming problem does not depend on the parameters of the supply-side, i.e. the demand relationship can be separately estimated, solution to the multiple state variables problem on the supply side is unaffected.

Consider a forward looking consumer of type i. Time is discrete and indexed by $t = 0, 1, 2...$, horizon is infinite. Suppose per-period utility function of a consumer type i is defined as in (1). Let Ω_t denote current service characteristics and all other factors that affect future service characteristics. The following assumption is typically made in the literature to reduce the dimensionality of the producer state space.

Assumption 2: *Vector Ω_t consists of a pair of current utility flows $(\delta_{ict}, \delta_{ist})$ and it evolves as a first-order Markov process $P(\Omega_{t+1}|\Omega_t) = P(\delta_{ict+1}, \delta_{ist+1}|\delta_{ict}, \delta_{ist})$.*

It is worth noting that assumption (2) implies bounded rationality in consumer behavior,

[6]The discussion in the next section follows traditional approach in the empirical IO literature by modeling consumer expectations with a reduced-form specification of beliefs.

when consumers do not explicitly consider producer optimization problem. In practice, this assumption allows for separate estimation of the demand relationship as long as the data is generated by a market equilibrium and the reduced form specification of the process $P(\delta_{ict+1}, \delta_{ist+1} | \delta_{ict}, \delta_{ist})$ is flexibly estimated.[7]

Consumer dynamic programming problem can be written recursively in the form of a Bellman equation

$$
\begin{aligned}
&V_i(\epsilon_{it}, \delta_{ict}, \delta_{ist}, a_{it-1}) = \\
&\max \left\{ \begin{array}{l} \epsilon_{iot} + \beta E \left[V_i \left(\epsilon_{it+1}, \delta_{ict+1}, \delta_{ist+1}, a_{it} = o \right) | \epsilon_{it}, \delta_{ict}, \delta_{ist}, a_{it-1} \right], \\ u_{ict}(\delta_{ict}, a_{it-1}) + \beta E \left[V_i \left(\epsilon_{it+1}, \delta_{ict+1}, \delta_{ist+1}, a_{it} = c \right) | \epsilon_{it}, \delta_{ict}, \delta_{ist}, a_{it-1} \right], \\ u_{ist}(\delta_{ist}, a_{it-1}) + \beta E \left[V_i \left(\epsilon_{it+1}, \delta_{ict+1}, \delta_{ist+1}, a_{it} = s \right) | \epsilon_{it}, \delta_{ict}, \delta_{ist}, a_{it-1} \right] \end{array} \right\}
\end{aligned} \tag{4}
$$

where $u_{ijt}(\delta_{ijt}, a_{it-1})$ are defined as in (1).

Under assumption (1) probability that consumer of type i chooses service j in the current period given last period choice k is

$$
\begin{aligned}
\Pr(k \to j) =& \Pr(V_i^{kj} + \epsilon_{ijt} \geq V_i^{kl} + \epsilon_{ilt}, \forall l \neq j\}) \\
=& \frac{\exp V_i^{kj}}{\sum_l \exp V_i^{kl}}
\end{aligned} \tag{5}
$$

where

$$
V_i^{kj}(\delta_{ict}, \delta_{ist}) = V(\delta_{ict}, \delta_{ist}, a_{it} = j, a_{it-1} = k)
$$

defines *"choice-specific"* value function *net* of current idiosyncratic preference draw ϵ_{ijt}. Note that unlike in the case of myopic consumers value of choosing outside alternative today is not zero as it encapsulates an option to subscribe to one of the paid television providers in the future. These choice-specific value functions can be computed using joint contraction mapping

$$
\left\{ \begin{array}{l} V_i^{oo} = \beta E \ln[\exp V_i^{oo} + \exp(V_i^{cc} - \eta_c) + \exp(V_i^{ss} - \eta_s)], \\ V_i^{cc} = \delta_{ict} + \beta E \ln[\exp V_i^{oo} + \exp V_i^{cc} + \exp(V_i^{ss} - \eta_s)], \\ V_i^{ss} = \delta_{ist} + \beta E \ln[\exp V_i^{oo} + \exp(V_i^{cc} - \eta_c) + \exp V_i^{ss}] \end{array} \right. \tag{6}
$$

where expectations are taken with respect to future values of $(\delta_{ict}, \delta_{ist})$ and $V_i^{kj} = -\eta_{ij} + V_i^{jj}$.

Then individual market shares for each type of consumers can be calculated similar to (2) and (3).

[7] In the modern literature on dynamic demand estimation, assumption (2) is often replaced with a much stronger assumption on the evolution of the entire market, which is captured by the "logit inclusive values" as in Melnikov (2001), Gowrisankaran and Rysman (2007), Schiraldi (2006).

5.3 Aggregate demand schedule

In the previous sections 5.1 and 5.2 individual shares for each consumer type i in period t were shown to depend on the distribution of the shares across available alternatives (s_{ict-1}, s_{ist-1}) in the previous period $t-1$ and current utility flows $(\delta_{ict}, \delta_{ist})$.

Let $s_{ict} = s_{ic}(\delta_{ict}, \delta_{ist}, s_{ict-1}, s_{ist-1})$ and $s_{ict} = s_{ic}(\delta_{ict}, \delta_{ist}, s_{ict-1}, s_{ist-1})$ describe the individual demand for consumer type i in period t given by (2) and (3) in case of myopic consumers or by their analogs in case of forward-looking consumer behavior.

In order to obtain aggregate market shares I need to integrate over the the distribution of consumer types, i.e.

$$s_{jt} = \int s_{ijt}(\delta_{ct}, \delta_{st}, s_{ict-1}, s_{ist-1}; \alpha_{ij}, \eta_{ij}, \tilde{\alpha}_i^p) dF(\alpha_{ij}, \eta_{ij}, \tilde{\alpha}_i^p | \mathbf{\Theta}) \quad (7)$$

where $(\delta_{ct}, \delta_{st})$ are defined in (1) and (s_{ict-1}, s_{ist-1}) is a distribution of the last period shares for each consumer type i.

If persistent consumer preferences are modeled by N draws from the distribution of consumer types $F(\alpha_{ij}, \eta_{ij}, \tilde{\alpha}_i^p | \mathbf{\Theta})$ aggregate demand relationship for cable service is given by

$$s_{ct} = \frac{1}{N} \sum_{i=1}^{N} s_{ic}(\delta_{ict}, \delta_{ist}, s_{ict-1}, s_{ist-1}; \Theta), \quad (8)$$

i.e. a simple frequency simulator.[8]

From (8) it is clear that current aggregate share depends on the entire vector of the last period market shares for each consumer type, i.e.

$$s_{jt} = s_j(s_{1ct-1}, s_{1st-1}, \ldots, s_{Nct-1}, s_{Nst-1}, \delta_{ct}, \delta_{st}; \Theta), \;\; j = c, s.$$

In general, if there are J products and N consumer types, the dimensionality of the vector defining current period market share is $(J-1) \times N$. Given that the vector enters state space of the producer dynamic programming problem, with an increase in the number of products and/or the number of simulated consumer types traditional solution of the supply-side problem becomes computationally infeasible.

Note that if one makes an assumption that instead of keeping track of individual consumers' past decisions a producer approximates the last period distribution of consumer types across alternatives with a finite number of moments of such distribution as in Krusell and Smith (1998) the number of relevant state variables reduces to $(J-1) \times \#MOMENTS$. However, with a large number of products full solution to the producer dynamic programming problem may still be extremely difficult.

In the next section, I discuss an alternative approach to estimating supply side parameters, which relies on the empirical version of the first-order conditions for the dynamic controls.

[8] In many cases, there are alternative simulation techniques (e.g. importance sampling) which might have better statistical properties.

5.4 Supply side

On the supply side of the paid television market there are two producers: cable and satellite providers of the paid television service. As in the previous sections, variables are indexed by a subscript $j \in \{o, c, s\}$.

Time is discrete and is indexed by $t = 0, 1, 2, \ldots$ I assume that each provider offers only a single tier.[9] The evolution of price, observed and unobserved quality offered by satellite provider is assumed to be exogenous. Therefore, the only strategic player in the market is cable companies.

Prior to making price, p_{ct}, and quality, q_{ct}, choice cable companies observe price, observed and unobserved (by econometricians) quality of satellite service. In addition, cable providers observe realizations of own scalar unobservable, ξ_{ct}, as well realizations of a vector of exogenous cost shifters Z_{ct}. The above is summarized by the following assumption.

Assumption 3: *In each market and each time period a cable company solves a single-agent decision problem after observing realizations of exogenous variables, $x_t = (p_{st}, q_{st}, \xi_{st}, \xi_{ct}, Z_{ct})$.*

As discussed in the section (5.3), a cable company faces demand relationship summarized by the following equation

$$D_{ct} = M_t \cdot \sum_{i=1}^{N} s_{ic}(p_{ct}, q_{ct}, x_t; s_{ict-1}, s_{ist-1}) \tag{9}$$

where M_t is the market size in period t and $s_{ic}(\cdot)$ is a function that computes current share of consumer type i.

Let $s_{it} = (s_{ict}, s_{ist})',\ i = 1, \ldots, N$ denote a pair of cable and satellite shares of consumer type i in period t. Define $s_t = (s_{it}, \ldots, s_{Nt})'$ be a $2N \times 1$ vector of all consumer types shares in period t.

Assume that market size is constant over time, i.e. $M_t = M,\ \forall t$ then per-period profit of a cable provider has the following parametric form

$$\Pi(s_{t-1}, x_t, p_{ct}, q_{ct}) = M \cdot \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct})(p_{ct} - c(q_{ct}, Z_{ct})), \tag{10}$$

where $c(q_{ct})$ is a marginal cost of providing quality q_{ct} per subscriber. Note that this specification implicitly assumes that the cost function is perfectly scalable in the number of subscribers.

[9]This assumption is mainly driven by the data limitations. Extension to multiple tiers possible and discussed later in the paper.

Assume that a cable company maximizes present discounted value of future cash flows over an infinite time horizon, i.e. the dynamic programming problem of the cable provider can be written as follows

$$W(s_{t-1}, x_t) = \max_{\{p_{ct}, q_{ct}\}} \left\{ E_t \sum_{t=0}^{\infty} \beta^t \Pi(s_{t-1}, x_t, p_{ct}, q_{ct}) \right\} \quad (11)$$
$$s.t.$$
$$s_t = s(s_{t-1}, x_t, p_{ct}, q_{ct})$$

where expectation is taken with respect to future values of the exogenous state variables x_t.

Recursive formulation of the producer dynamic programming problem is presented by the following Bellman equation

$$W(s_{t-1}, x_t) = \max_{p_{ct}, q_{ct}} \{\Pi(s_{t-1}, x_t, p_{ct}, q_{ct}) + \beta E_t[W(s_t, x_{t+1})]\} \quad (12)$$
$$s.t.$$
$$s_{ict} = s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}), \forall i = 1, \ldots, N;$$
$$s_{ist} = s_{is}(s_{it-1}, x_t, p_{ct}, q_{ct}), \forall i = 1, \ldots, N;$$

where $E_t[\cdot]$ stands for expectation over future values of x_t conditional on the information set in period t.

In the beginning of each period, cable company observes realizations of the exogenous stochastic variables x_t and chooses (p_{ct}, q_{ct}). Let $p(s_{t-1}, x_t)$ and $q(s_{t-1}, x_t)$ denote optimal policy functions. Note that conditional on the current period values of x_t transitions of individual market shares are deterministic, i.e. by choosing (p_{ct}, q_{ct}) the producer knows next period values of the state variables s_t. Therefore, the only source of uncertainty is embedded into the x_t process.

First order conditions of the right-hand side of (12) with respect to dynamic controls (p_{ct}, q_{ct}) are

$$\frac{\partial \Pi(s_{t-1}, x_t, p_{ct}, q_{ct})}{\partial p_{ct}} + \beta \left[\frac{\partial s_t}{\partial p_{ct}}\right]^T E_t \left[\frac{\partial W(s_t, x_{t+1})}{\partial s_t}\right] = 0 \quad (13)$$

$$\frac{\partial \Pi(s_{t-1}, x_t, p_{ct}, q_{ct})}{\partial q_{ct}} + \beta \left[\frac{\partial s_t}{\partial q_{ct}}\right]^T E_t \left[\frac{\partial W(s_t, x_{t+1})}{\partial s_t}\right] = 0 \quad (14)$$

where

$$\left[\frac{\partial s_t}{\partial p_{ct}}\right]^T = \begin{pmatrix} \frac{\partial s_{1c}(\cdot)}{\partial p_{ct}} & \frac{\partial s_{1s}(\cdot)}{\partial p_{ct}} & \ldots & \frac{\partial s_{Nc}(\cdot)}{\partial p_{ct}} & \frac{\partial s_{Ns}(\cdot)}{\partial p_{ct}} \end{pmatrix}$$
$$\left[\frac{\partial s_t}{\partial q_{ct}}\right]^T = \begin{pmatrix} \frac{\partial s_{1c}(\cdot)}{\partial q_{ct}} & \frac{\partial s_{1s}(\cdot)}{\partial q_{ct}} & \ldots & \frac{\partial s_{Nc}(\cdot)}{\partial q_{ct}} & \frac{\partial s_{Ns}(\cdot)}{\partial q_{ct}} \end{pmatrix}$$

and

$$\left[\frac{\partial W(s_t, x_{t+1})}{\partial s_t}\right] = \begin{pmatrix} \dfrac{\partial W(s_t, x_{t+1})}{\partial s_{1ct}} \\ \dfrac{\partial W(s_t, x_{t+1})}{\partial s_{1st}} \\ \vdots \\ \dfrac{\partial W(s_t, x_{t+1})}{\partial s_{Nct}} \\ \dfrac{\partial W(s_t, x_{t+1})}{\partial s_{Nst}} \end{pmatrix}$$

Consider i^{th} elements of the vector $\left[\frac{\partial s_t}{\partial p_{ct}}\right]$. The implications of the assumption 1 are as follows

$$\begin{aligned} \frac{\partial s_{ic}(\cdot)}{\partial p_{ct}} =& s_{ict-1}\frac{\partial Pr_i(c \to c)}{\partial p_{ct}} + s_{ist-1}\frac{\partial Pr_i(s \to c)}{\partial p_{ct}} + (1 - s_{ict-1} - s_{ist-1})\frac{\partial Pr_i(o \to c)}{\partial p_{ct}} \\ =& \alpha_i^p \begin{pmatrix} s_{ict-1} \cdot \big(Pr_i(c \to c) - Pr_i^2(c \to c)\big) \\ + s_{ist-1} \cdot \big(Pr_i(s \to c) - Pr_i^2(s \to c)\big) \\ + (1 - s_{ict-1} - s_{ist-1}) \cdot \big(Pr_i(o \to c) - Pr_i^2(o \to c)\big) \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \frac{\partial s_{is}(\cdot)}{\partial p_{ct}} =& s_{ist-1}\frac{\partial Pr_i(s \to s)}{\partial p_{ct}} + s_{ict-1}\frac{\partial Pr_i(c \to s)}{\partial p_{ct}} + (1 - s_{ict-1} - s_{ist-1})\frac{\partial Pr_i(o \to s)}{\partial p_{ct}} \\ =& -\alpha_i^p \begin{pmatrix} s_{ist-1} \cdot Pr_i(s \to s) \cdot Pr_i(s \to c) \\ + s_{ict-1} \cdot Pr_i(c \to s) \cdot Pr_i(c \to c) \\ + (1 - s_{ict-1} - s_{ist-1}) \cdot Pr_i(o \to s) \cdot Pr_i(o \to c) \end{pmatrix} \end{aligned}$$

Similarly, the i^{th} element of the vector $\left[\dfrac{\partial s_t}{\partial q_{ct}}\right]$ is given by

$$\frac{\partial s_{ic}(\cdot)}{\partial q_{ct}} = \alpha_i^q \begin{pmatrix} s_{ict-1} \cdot \big(Pr_i(c \to c) - Pr_i^2(c \to c)\big) \\ + s_{ist-1} \cdot \big(Pr_i(s \to c) - Pr_i^2(s \to c)\big) \\ + (1 - s_{ict-1} - s_{ist-1}) \cdot \big(Pr_i(o \to c) - Pr_i^2(o \to c)\big) \end{pmatrix}$$

$$\frac{\partial s_{is}(\cdot)}{\partial q_{ct}} = -\alpha_i^q \begin{pmatrix} s_{ist-1} \cdot Pr_i(s \to s) \cdot Pr_i(s \to c) \\ + s_{ict-1} \cdot Pr_i(c \to s) \cdot Pr_i(c \to c) \\ + (1 - s_{ict-1} - s_{ist-1}) \cdot Pr_i(o \to s) \cdot Pr_i(o \to c) \end{pmatrix}$$

Finally, derivatives of the current period reward function with respect to p_{ct} are given by the following formulas

$$\begin{aligned} \frac{\partial \Pi(s_{t-1}, x_t, p_{ct}, q_{ct})}{\partial p_{ct}} &= \sum_{i=1}^{N} s_{ict} + \sum_{i=1}^{N} \frac{\partial s_{ic}}{\partial p_{ct}} (p_{ct} - c(q_{ct})) \\ &= s_{ct} + (p_{ct} - c(q_{ct})) \sum_{i=1}^{N} \alpha_i^p \begin{pmatrix} s_{ict-1}(Pr_i(c \to c) - Pr_i^2(c \to c)) + \\ s_{ist-1}(Pr_i(s \to c) - Pr_i^2(s \to c)) + \\ (1 - s_{ict-1} - s_{ist-1})(Pr_i(o \to c) - Pr_i^2(o \to c)) \end{pmatrix} \end{aligned}$$

And the derivatives of the reward function with respect to q_{ct} are

$$\begin{aligned} \frac{\partial \Pi(s_{t-1}, x_t, p_{ct}, q_{ct})}{\partial q_{ct}} = &- s_{ct} \frac{\partial c(q_{ct})}{\partial q_{ct}} + \\ &(p_{ct} - c(q_{ct})) \sum_{i=1}^{N} \alpha_i^q \begin{pmatrix} s_{ict-1}(Pr_i(c \to c) - Pr_i^2(c \to c)) + \\ s_{ist-1}(Pr_i(s \to c) - Pr_i^2(s \to c)) + \\ (1 - s_{ict-1} - s_{ist-1})(Pr_i(o \to c) - Pr_i^2(o \to c)) \end{pmatrix} \end{aligned}$$

In the next section, I discuss an estimation strategy that relies on the empirical version of the first order-conditions (13) and (14). The approach is similar to the generalized method of moments for non-linear rational expectations model proposed by Hansen and Singleton (1982). There is, however, a significant difference stemming from the fact that dynamic controls p_{ct} and q_{ct} shift a large number of state variables at the same time. Under these conditions it is very hard to derive empirical Euler equations that are expressed in terms of the model primitives, i.e. per-period payoff functions and derivatives of the transition functions. Hence, to estimate from the first-order conditions I need to know the derivatives of the value function with respect to the vector of state variables. To overcome the problem, I use forward simulation technique similar to Hotz et al. (1994) and Bajari et al. (2007).

6 Estimation strategy

In this section I discuss estimation algorithms used to estimate demand and supply side parameters and the next section provides identification arguments and defines sets of instrumental variables for estimation.

6.1 Demand side

Estimation of the demand side parameters follows approach originally suggested by Berry et al. (1995) and the literature that follows. In particular, aggregate market shares predicted by the model (given parameter vector Θ) computed as in (8) can be written as functions of the population mean utility flows $(\delta_{ct}, \delta_{st})$

$$s_c(\delta_{ct}, \delta_{st}, p_{ct}, p_{st}, s_{it-1}; \Theta) = \frac{1}{N} \sum_{i=1}^{N} s_{ic}(\delta_{ict}(\delta_{ct}, p_{ct}), \delta_{ist}(\delta_{st}, p_{st}), s_{ict-1}, s_{ist-1}; \Theta)$$
$$s_s(\delta_{ct}, \delta_{st}, p_{ct}, p_{st}, s_{it-1}; \Theta) = \frac{1}{N} \sum_{i=1}^{N} s_{is}(\delta_{ict}(\delta_{ct}, p_{ct}), \delta_{ist}(\delta_{ct}, p_{st}), s_{ict-1}, s_{ist-1}; \Theta)$$

The key idea behind estimation algorithm is to solve for the unknown mean population utility flows such that model predictions match observed aggregate market shares, i.e. to find such $(\hat{\delta}_{ct}, \hat{\delta}_{st})$ that solve

$$\begin{cases} s_{ct} = s_c(\hat{\delta}_{ct}, \hat{\delta}_{st}, p_{ct}, p_{st}, s_{it-1}; \Theta), \\ s_{st} = s_s(\hat{\delta}_{ct}, \hat{\delta}_{st}, p_{ct}, p_{st}, s_{it-1}; \Theta) \end{cases}$$

To do this define the following fixed point equations

$$\begin{cases} \delta_{ct}^{n+1} = \delta_{ct}^{n} + (\ln(s_{ct}) - \ln(s_c(\delta_{ct}^{n}, \delta_{st}^{n}, p_{ct}, p_{st}, s_{it-1}; \Theta)), \\ \delta_{st}^{n+1} = \delta_{st}^{n} + (\ln(s_{st}) - \ln(s_s(\delta_{ct}^{n}, \delta_{st}^{n}, p_{ct}, p_{st}, s_{it-1}; \Theta)), \end{cases} \tag{15}$$

where n and $n+1$ denote current and next iteration values of the unknown variables.

Let $(\hat{\delta}_{ct}, \hat{\delta}_{st})$ denote "inverted-out" values of the population mean flow utilities given parameter vector Θ. Under scalar unobservable assumption and per-period utility specification (1) for any values of the demand-side parameters I can recover $(\hat{\xi}_{ct}, \hat{\xi}_{st})$.[10]

[10]The fixed-point equations (15) could be defined directly in terms of (ξ_{ct}, ξ_{st}). However, from the computational perspective this is not desirable because when the per-period utility function is linear in parameters there is a closed form solution that reduces the number of parameters to search for in a non-linear optimization routine.

To estimate demand-side parameters I assume that there exists a vector of instruments Z_t that is orthogonal to the current period *innovations* to the unobserved service characteristics (ξ_{ct}, ξ_{st}) and is correlated with the endogenous variables p_{ct} and q_{ct}. In particular, I make the following assumption

Assumption 4: *Evolution of the unobserved service characteristics (ξ_{ct}, ξ_{st}) satisfies the following equations*

$$\xi_{ct} = \xi_{ct-1} + v_{ct},$$
$$\xi_{st} = \xi_{st-1} + v_{st}$$

and there exists a vector of instruments Z_t such that

$$E\left[\begin{matrix} v_{ct} \\ v_{st} \end{matrix}\middle| Z_t\right] = 0$$

Particular set of instrumental variables that include supply-side "cost-shifters" is discussed in section 7 below. At this point it is worth noting that identification of the switching costs would require finding exogenous variables that affected last period consumer decisions.

In case of the forward-looking consumers, one can use similar estimation strategy except for one extra computational step that solves for the consumer value functions. In particular, for any given value of the parameter vector Θ one has to assume initial parameters for the consumer beliefs specified by the process $P(\delta_{ict+1}, \delta_{ist+1}|\delta_{ict}, \delta_{ist})$. Given that, solve for the consumer value functions using joint contraction mapping (6) and invert out population mean utility flows $(\delta_{ct}, \delta_{st})$ by matching market shares predictions to the observed data. Note, however, that for the first several iterations the evolution of the recovered $(\delta_{ict}, \delta_{ist})$ quantities generally will not be consistent with the initial guess of the parameters for the consumer beliefs (about this evolution). Therefore, the reduced form specification for the consumer beliefs needs to be updated using the realizations of $(\delta_{ict}, \delta_{ist})$ and the dynamic programming problem has to be re-solved again. Hence, the steps above need to be repeated up until convergence, i.e. when consumer problem is solved using beliefs that are consistent with the actual evolution of the individual flow utilities that, in turn, were inverted out using the solution to the consumer dynamic programming problem.

One last issue regarding estimation that needs to be discussed is the initial conditions problem. In case of the paid-television industry the data from the early years when the cable television was first introduced is missing. As I describe in section 4, I observe cable companies market shares, prices and quality only since 1992.

Note that the evolution of aggregate market shares as defined in (7) and its simulated counterpart (8) depends on the initial distribution of the consumer types across alternative services, i.e. $(s_{i,1992}), \forall i$. In order to obtain a guess about the initial distribution I assume that prior to 1992 there were no paid television. Further, I assume that in 1992 consumers face only two alternatives, i.e. cable and over-the-air television. Given that DBS companies effectively entered the market in 1993 I add the satellite alternative in 1993. Given that the initial distribution of consumer types was approximated in this way, I use data from 1992 till 1996 only to simulate initial conditions for the years 1997-2002.[11] The hope is that by 1997 variation in the cable and satellite prices and qualities alleviates the impact of the initial conditions assumption and the distribution of the consumer types across alternatives is close to the true one.

6.2 Supply side

Parameters of the supply side can be estimated using generalized instrumental variables approach to the nonlinear rational expectations model developed in Hansen and Singleton (1982). In particular, define the following functions

$$g_p(s_{t-1}, x_t, p_{ct}, q_{ct}, s_t, x_{t+1}) = \frac{\partial \Pi(s_{t-1}, x_t, p_{ct}, q_{ct})}{\partial p_{ct}} + \beta \left[\frac{\partial s_t}{\partial p_{ct}}\right]^T E_t \left[\frac{\partial W(s_t, x_{t+1})}{\partial s_t}\right] \tag{16}$$

$$g_q(s_{t-1}, x_t, p_{ct}, q_{ct}, s_t, x_{t+1}) = \frac{\partial \Pi(s_{t-1}, x_t, p_{ct}, q_{ct})}{\partial q_{ct}} + \beta \left[\frac{\partial s_t}{\partial q_{ct}}\right]^T E_t \left[\frac{\partial W(s_t, x_{t+1})}{\partial s_t}\right] \tag{17}$$

Rational expectations assumption suggests that the following must be true

$$E_t[g_p(s_{t-1}, x_t, p_{ct}, q_{ct})] = 0 \tag{18}$$

$$E_t[g_q(s_{t-1}, x_t, p_{ct}, q_{ct})] = 0 \tag{19}$$

when evaluated at the observed (supposedly optimal) choices of the dynamic controls (p_{ct}, q_{ct}).[12]

There are several issues to consider. First, to construct empirical versions of $g_p(\cdot)$ and $g_q(\cdot)$ I need to know vector $(s_{t-1}, x_t, p_{ct}, q_{ct})$. One possibility to proceed is to use the estimates of the consumer types distribution, $\hat{s}_t$, and unobserved service characteristics $(\hat{\xi}_{ct}, \hat{\xi}_{st})$ obtained from the demand side.[13]

[11] There is another reason to omit data prior to 1997. Satellite penetration rates by DMA are available only from 1997. To backcast the satellite shares by market in 1993-1996 I used national-level dynamics in satellite shares and their values in 1997.

[12] Note that if policy functions are known exactly the conditions above must be equal to zero when evaluated at observed price and quality choices, i.e. the model "overfits" the data. In estimation, however, whenever possible, I use observed realizations of the stochastic variables in the vector x_t and simulate their evolution only after the terminal period in the data. This, in turn, generates deviations from the first-order conditions. In the new version of the paper I propose a way to introduce supply-side unobservables in a structural way.

[13] Note that computation of the standard errors for the supply side parameters in this case should account for the "first-stage" demand side estimation.

Second, constructing GMM estimator using (18) and (19) requires knowledge of a vector of partial derivatives of the value function with respect to the state vector, i.e. $\left[\dfrac{\partial W(s_t, x_{t+1})}{\partial s_t}\right]$. This task is more challenging as it is not possible to express the derivatives in terms of the model primitives as it is often done in the literature on empirical Euler equations. To overcome the problem I suggest using simulation technique in the spirit of Hotz et al. (1994) and Bajari et al. (2007).

First, I estimate policy functions $p(s_{t-1}, x_t)$ and $q(s_{t-1}, x_t)$ using the data, estimates of the distribution of consumer shares, and estimates of the evolution of exogenous variables x_t obtained from the demand side.

Then I simulate forward NS possible paths (each of T periods length) of exogenous variables, resulting policies and implied transitions of the endogenous variables. By computing sequences of the reward functions for each period and averaging over NS simulations I can get an approximation to the continuation value for any given starting point (s_{t-1}, x_t).

In particular, let $\hat{p}(s_{t-1}, x_t; \theta)$ and $\hat{q}(s_{t-1}, x_t; \theta)$ denote parametric estimates of the policy functions. Then an approximation to the value function from following policies $\hat{p}(s_{t-1}, x_t; \theta)$ and $\hat{q}(s_{t-1}, x_t; \theta)$ is given by

$$\hat{W}(s_{t-1}, x_t) = \frac{1}{NS} \sum_{n=1}^{NS} \sum_{t=0}^{T} \beta^t \Pi(s_{t-1}, x_t, \hat{p}_{ct}(s_{t-1}, x_t), \hat{q}_{ct}(s_{t-1}, x_t)) \tag{20}$$

and implied transition of the endogenous variables $s_t = s(s_{t-1}, \hat{p}_{ct}(s_{t-1}, x_t), \hat{p}_{ct}(s_{t-1}, x_t))$.

Consider two perturbations to the vectors of states $s_{t-1} + \epsilon \iota_j$ and $s_{t-1} - \epsilon \iota_j$ where ι_j is a vector of dimension $\dim(s_{t-1}) \times 1$ with 1 in position j and zeros everywhere else. Then the derivative of value function with respect to a state variable j can be approximated as

$$\frac{\partial \hat{W}(s_{t-1}, x_t)}{\partial s_{jt-1}} = \frac{\hat{W}(s_{t-1} + \epsilon \iota_j, x_t) - \hat{W}(s_{t-1} - \epsilon \iota_j, x_t)}{2\epsilon} \tag{21}$$

In a similar way I can obtain approximations to derivatives with respect to each element of s_{t-1}. Note that this is still a very computationally intense procedure because I need to simulate $\dim(s_{t-1})$ derivatives for *each* point in the data.

As it is pointed out in Bajari et al. (2007), significant saving in terms of computation time comes from linearity of the reward function in the parameters of interest. In particular, if per-period reward function can be written as

$$\Pi(s_{t-1}, x_t, p_{ct}, q_{ct}; \theta) = \mathbf{\Psi}(s_{t-1}, x_t, p_{ct}, q_{ct}) \cdot \theta$$

where $\Psi(s_{t-1}, x_t, p_{ct}, q_{ct})$ is an $M-$dimensional vector of basis functions, the computationally intensive approximation of partial derivatives of the value function needs to be performed only once. Then for every trial value of the parameter vector I can use the same basis functions previously stored in the memory.

In particular, suppose that the cost of providing quality q_{ct} per subscriber has the following parametric representation

$$c(q_{ct}) = \theta_0 + \theta_1 q_{ct} + \theta_2 q_{ct}^2 + \theta_3 Z_{1ct} + \cdots + \theta_{K+2} Z_{Kct} \tag{22}$$

where θ_3 through θ_{K+2} are parameters for K cost shifters, $Z_{1ct}, \ldots, Z_{Kct}$.

Then I can write

$$\begin{aligned}
\Pi(s_{t-1}, x_t, p_{ct}, q_{ct}) = & \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) p_{ct} - \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) \cdot \theta_0 \\
& - \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) q_{ct} \cdot \theta_1 - \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) q_{ct}^2 \cdot \theta_2 \\
& - \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) \cdot Z_{1ct}\theta_3 \cdots - \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) \cdot Z_{Kct}\theta_{K+2}
\end{aligned}$$

Define

$$\begin{aligned}
\Psi_1(s_{t-1}, x_t, p_{ct}, q_{ct}) &= \frac{1}{NS} \sum_{n=1}^{NS} \sum_{t=0}^{T} \beta^t \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) p_{ct} \\
\Psi_2(s_{t-1}, x_t, p_{ct}, q_{ct}) &= \frac{1}{NS} \sum_{n=1}^{NS} \sum_{t=0}^{T} \beta^t \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) \\
\Psi_3(s_{t-1}, x_t, p_{ct}, q_{ct}) &= \frac{1}{NS} \sum_{n=1}^{NS} \sum_{t=0}^{T} \beta^t \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) q_{ct} \\
\Psi_4(s_{t-1}, x_t, p_{ct}, q_{ct}) &= \frac{1}{NS} \sum_{n=1}^{NS} \sum_{t=0}^{T} \beta^t \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) q_{ct}^2 \\
\Psi_5(s_{t-1}, x_t, p_{ct}, q_{ct}) &= \frac{1}{NS} \sum_{n=1}^{NS} \sum_{t=0}^{T} \beta^t \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) Z_{1ct} \\
\cdots \\
\Psi_{K+4}(s_{t-1}, x_t, p_{ct}, q_{ct}) &= \frac{1}{NS} \sum_{n=1}^{NS} \sum_{t=0}^{T} \beta^t \sum_{i=1}^{N} s_{ic}(s_{it-1}, x_t, p_{ct}, q_{ct}) Z_{Kct}
\end{aligned}$$

Let $\psi_k^j = \dfrac{\Psi_k(s_{t-1} + \epsilon\iota_j, x_t, p_{ct}, q_{ct}) - \Psi_k(s_{t-1} - \epsilon\iota_j, x_t, p_{ct}, q_{ct})}{2\epsilon}, k = 1, \ldots, K+4$. Then approximation to the derivatives of value function can be computed as follows

$$\frac{\partial \hat{W}(s_{t-1}, x_t)}{\partial s_{jt-1}} = \psi_1^j - \psi_2^j \theta_0 - \psi_3^j \theta_1 \cdots - \psi_{K+4}^j \theta_{K+2}$$

where the values of the basis functions are computed once and stored in memory. As long as all necessary basis functions are computed and stored in the computer memory they need not be re-computed for each trial value of the supply side parameters.

7 Instruments and identification

According to the assumption 3, cable companies observe realizations of $(p_{st}, q_{st}, \xi_{st}, \xi_{ct}, Z_{ct})$ prior to making their price and quality choices. Therefore, price and quality variables are clearly correlated with both the level of unobserved service characteristics and the current period innovations to them. In order to find instruments for the demand-side estimation I use the following arguments.

First, possible instruments for price and quality levels of cable providers are average prices and quality levels of other cable systems that belong to the same multiple-system-operator (MSO). These variables must be uncorrelated with the unobserved local market service characteristics, ξ's, but should be reasonable proxies for the price and quality levels offered by the local cable system. Correlation in prices and quality levels across systems occurs because the owner of several cable systems typically negotiates programming fees and other contract arrangements with programming networks on behalf of all of its members simultaneously. In turn, correlation in the marginal costs of systems within the same MSO justifies correlation in their price and quality levels. For the instruments to be valid, one must ensure that the unobserved demand shocks, ξ's are not correlated across the systems. It is less obvious because MSO typically own geographically concentrated firms. If unobserved demand shocks are closely correlated across different cable markets, the validity of these instruments may be questionable.

Second, different MSO's have different bargaining power in negotiations with programming networks. It is conceivable that larger MSO's with bigger number of subscribers to their cable systems have stronger bargaining position. Hence, I used the number of MSO subscribers as another instrument that shifts costs of all its members.

Third, programming networks often sell bundles consisting of several channels. Ability to purchase such bundles depends on the capacity level (in terms of the maximum number of channels physically possible to transmit through the cable system). Hence, average capacity level within MSO should be correlated with the ability of their member-systems to get lower rates. By the same logic, I used own capacity level as another instrumental variable. It is possible that systems with more favorable innovations into their ξ_{ct} would have higher marginal profitability of capital and, hence, would invest more actively into own system capacity. However, it is less likely that the system's capacity level would immediately respond to the current period innovation.

Fourth, total length of own coaxial lines of the local cable systems is a proxy for the differences in maintenance costs incurred by the systems in areas with different densities of houses.

Fifth, by exploiting peculiarities of the paid-television industry where satellite companies set their prices and quality at the national level I can use p_{st} and q_{st} as instruments for local cable price and quality variables. The argument here is that DBS prices and quality are not likely to respond to changes in the local demand unobservables (ξ_{ct}, ξ_{st}), while cable companies choose price and quality locally after observing realizations of p_{st} and q_{st}.[14]

So far I described instruments used to solve the problem of price and quality endogeneity. Given that the model has extra parameters for the distribution of switching costs, it is necessary to discuss the instruments that help to identify them. Non-trivial switching costs generate state dependence in consumer utility. Loosely speaking, switching costs are "coefficients" on the lagged market shares if we were running linear regression. Exogenous variables that were relevant for the previous period consumer choices should serve as informative additional instruments to identify consumer switching costs. Therefore, in addition to the current period values of the instrumental variables discussed above I included their lagged values.

Finally, I used another set of IV's that should enhance identification of the switching costs parameters.[15] Moving decisions are likely to exogenously "reset" last period consumer state more often in the regions with high population mobility. To proxy for population mobility I used the number of housing permits issued at the state level.

Up until recently, in order to install satellite dish in a multy-story buildings consumers have to obtain a permission from the owner, which complicated usage of the satellite television in regions with a large number of multiunit buildings. This motivates including variables like percentage of dwelling units in multiunit structures of 5 or more units.[16]

Another complication that many satellite subscribers face is related to the "southern exposure" issue, i.e. the necessity to locate the receiver (dish) in a place that guarantees open access to the orbital satellite. To control for the geographic location of a particular market I used its longitude and latitude as two extra instruments.

8 Estimation results

In this section I present preliminary results for the demand and supply side estimation. Demand side parameters were estimated using GMM based on a sample of 564 local cable

[14] A problem may arise if the unobserved service characteristics or, more generally, demand-side unobservables (ξ_{ct}, ξ_{st}) are correlated across markets and are taken into considerations by the satellite providers.

[15] Note that some of these IV's are not explicitly in the model. Some of them are easy to explicitly incorporate into a myopic consumer model than into the dynamic one. In the future versions of the paper I devote more space to discussing such possibilities.

[16] Also, it is conceivable that serving multiunit building is cheaper for a cable company than serving an equivalent number of separate houses.

markets in 1992-2002. Due to the data limitations discussed in section 6.1, to form the empirical moment conditions I used data from 1997-2002.

Given parameter estimates of the demand side, I augmented observed data with the estimates of the unobserved service characteristics of cable and satellite services and estimated parameters of the cost function for the supply-side model.[17]

8.1 Demand side

Demand side structural parameters were estimated using assumption of consumer myopia. To control for persistent consumer heterogeneity I used 30 consumer types whose preferences are represented by the iid draws from a normal distribution with a diagonal variance-covariance matrix. Table 3 presents results of the random coefficients model.

Table 3: 30-types model estimation results, GMM

Variable	1^{st} stage	2^{nd} stage
Switching cost (cab)	0.949	1.151
(s.e.)	(0.521)	(0.591)
σ_{η_c}	9.23e-05	8.03e-05
(s.e.)	(40.98)	(48.16)
Switching cost (sat)	2.238	1.844
(s.e.)	(0.370)	(0.442)
σ_{η_s}	0.0001	7.95e-05
(s.e.)	(29.44)	(32.75)
σ_{const_c}	4.03e-05	3.33e-05
(s.e.)	(38.42)	(40.49)
σ_{const_s}	9.02e-05	9.66e-05
(s.e.)	(2.737)	(3.442)
Price coefficient	-0.047	-0.087
(s.e.)	(0.008)	(0.010)
σ_{η_p}	1.38e-06	1.42e-06
(s.e.)	(0.370)	(0.435)
Quality coefficient	0.072	0.051
(s.e.)	(0.011)	(0.011))
F-value:		

Note that none of the parameters of consumer heterogeneity were precisely estimated. Point estimates of the standard deviations for the switching costs and constant term pa-

[17]Note that standard errors for the supply side parameters do not account for the first stage demand estimation. Necessary adjustment will be made in the next version of the paper.

rameters are fairly close to zero. The same is true for the standard deviation of the price coefficient. Standard errors are huge for all of the standard deviations.

On the other hand, the levels of switching costs as well as price and quality coefficients are estimated precisely. Note that since moment conditions were constructed using innovations into the unobserved service characteristics (ξ_{ct}, ξ_{st}), i.e. the latter ones were first-differenced, any constant terms in the *mean* utility specification are not separately identified from the means of the unobservables.

Given that the data fails to reject a representative consumer model, I re-estimated the demand side model under assumption of one consumer type. The results are presented in table 4.

Table 4: One-type model estimation results, GMM

Variable	1^{st} stage	2^{nd} stage	Cont. updated
Switching cost (cab)	0.949	1.036	1.171
(s.e.)	(0.488)	(0.560)	(0.571)
Switching cost (sat)	2.238	1.844	1.868
(s.e.)	(0.230)	(0.294)	(0.301)
Price coefficient	-0.047	-0.087	-0.094
(s.e.)	(0.005)	(0.005)	(0.006)
Quality coefficient	0.072	0.051	0.046
(s.e.)	(0.005)	(0.005)	(0.005)
F-value:	67.8	452.8	339.9

The level of consumer switching costs for cable and satellite providers was estimated at \$149 and \$238 respectively.[18]

8.2 Supply side

In order to augment the data for estimation of the parameters of the cable companies cost function I use fitted values of the unobserved quantities $(\hat{\xi}_{ct}, \hat{\xi}_{st})$ obtained from the demand side estimates. Since parameters of the consumer heterogeneity have point estimates close to zero and are not statistically significant at any reasonable significance level, the distribution of consumer heterogeneity collapses to a simple mean market shares for cable and satellite.

As discussed in section 6.2, to estimate parameters of the cable companies cost function

[18] These estimates are slightly higher than the ones obtained from the dynamic demand specification. In another paper Shcherbakov (2008) I find that cable and satellite switching costs were approximately \$109 and \$186. The difference in the estimates can be attributed to both the differences in the model framework and the differences in the data used for estimation.

I first estimated policy functions

$$p_{ct} = p(s_{t-1}, x_t; \theta)$$
$$q_{ct} = q(s_{t-1}, x_t; \theta)$$

where x_t includes satellite price and quality, average MSO price and quality, own and MSO capacity levels, miles of coaxial cable (in thousands), the number of MSO subscribers (in thousands), last period market shares, and the estimates of the unobserved cable and satellite service characteristics. The estimation results can be found in Appendix 1.

I used the estimated policy functions to simulate 1,000 sequences of 500 periods each to estimate basis functions for the numerical derivatives approximation. For the preliminary estimation I assumed that all of the exogenous state variables with the exception of unobserved service characteristics (ξ_{ct}, ξ_{st}) evolve deterministically till the terminal period in the data and then stay constant forever. This is a very strong assumption and at the moment I am working on a long-run model of the cable firms behavior that would allow for investment and stochastic evolution of the cost shifters.

The evolution of the exogenous variables (ξ_{ct}, ξ_{st}) is assumed to satisfy a random walk assumption with simulated innovations drawn from normal distribution with the variance equal to the empirical variance of the observed innovations. Conditional on the current simulated realizations of the cable and satellite unobservables I predict policies $(\hat{p}_{ct}, \hat{q}_{ct})$ and update current values of the endogenous variables (s_{ct}, s_{st}).

To estimate parameters of the cable companies cost function I use moment conditions specified in (18) and (19). The set of instruments includes own price and quality levels, price and quality of satellite companies, and the entire set of cost shifters. Preliminary estimation results are listed in table 5.

According to the estimates of the cable cost function, per-subscriber costs of providing service is increasing in the quality level at a slightly increasing rate.

Most of the coefficients on the cost shifters are statistically significant and have expected signs. In particular, the cost is increasing in average MSO price and decreasing in average MSO quality, which is consistent with the intuition behind this cost shifter. Cost of providing television service is decreasing in own capacity level. This suggests non-trivial gains to high capacity levels that allow for cost reduction when purchasing bundles of programming networks rather than individual channels.

Coefficient on the own length of coaxial lines is positive though not statistically significant. Unexpectedly, average MSO capacity level seems to have significant positive effect on the own cost structure of cable providers.

Finally, the average number of MSO subscribers tends to reduce costs of the cable companies.

Current estimates of the cable cost function suggest that the average level of costs was about \$2.19 (median \$2.12) per subscriber with the average price-cost margins of about

Table 5: Supply-side model, GMM

	Dynamic producer		Myopic producer	
Variable	First stage	Second stage	First stage	Second stage
Const.	-14.1270	-14.9800	-18.6780	-19.0230
(*s.e.*)	(0.8157)	(0.8011)	(1.0816)	(1.0865)
qc	0.4884	0.4916	0.4770	0.4884
(*s.e.*)	(0.0006)	(0.0001)	(0.0011)	(0.0003)
qc^2	0.0021	0.0005	0.0042	0.0011
(*s.e.*)	(0.0001)	(0.0000)	(0.0002)	(0.0000)
$MSO(pc)$	1.0514	1.0412	1.0002	0.9810
(*s.e.*)	(0.0216)	(0.0213)	(0.0312)	(0.0317)
$MSO(qc)$	-0.5124	-0.6250	-0.8816	-1.0590
(*s.e.*)	(0.0966)	(0.0952)	(0.1385)	(0.1390)
CAP	-0.0992	-0.1006	-0.1110	-0.1193
(*s.e.*)	(0.0084)	(0.0083)	(0.0134)	(0.0137)
$MCOAX$	0.4738	0.0450	-0.1892	-0.0125
(*s.e.*)	(0.3378)	(0.3573)	(0.7165)	(0.7198)
$MSO(CAP)$	0.0220	0.0691	0.0790	0.1109
(*s.e.*)	(0.0170)	(0.0165)	(0.0233)	(0.0236)
$MSO(SUB)$	-0.1019	-0.1085	-0.1806	-0.2094
(*s.e.*)	(0.0176)	(0.0172)	(0.0269)	(0.0278)
F-value	365.9	403.5	482.6	439.66

$16.52 (median $16.42). Interestingly, in some cases the estimates suggest negative price-cost margins of as low as $-3.04 per subscriber. Unless the negative estimates is an artifact of incorrectly specified supply and demand models, this suggests importance of dynamic price and quality policies, when producers invest heavily into the customer base by using introductory pricing in the initial periods.

The myopic model estimates (i.e. when discount factor on the continuation value was set to zero), suggests that the average level of costs is -$4.14 (median -$3.77). The price-cost margin is estimated at the mean level of $22.85 (median $22.65). Hence, the estimates from the myopic model are less plausible given the data on average prices in cable industry.

9 Counterfactual simulations

In this section I provide two counterfactual simulations that rely on the preliminary demand and supply side estimates. The first scenario evaluates the effect of switching costs on the cable television prices under assumption that prices and quality levels of the satellite company as well as own cost structure of the cable providers do not change. This outcome corresponds to a static duopoly model. The second scenario calculates new prices under no-satellite-no-switching-costs assumption, i.e. a static monopoly model.

Before discussing the results of the counterfactual simulations it is worth noting one peculiarity of a representative consumer model with linear consumer utility function. Consider a general formulation of the producer dynamic programming problem where demand schedule is given by myopic consumers who nevertheless face significant switching costs.[19]

$$W(s_{ct-1}, s_{st-1}, x_{ct}) = \max_{p_{ct}, q_{ct}} \{s(s_{ct-1}, s_{st-1}, x_{ct}, p_{ct}, q_{ct})(p_{ct} - c(q_{ct}, Z_{ct})) + \beta EW(s_{ct}, s_{st}, x_{ct+1})\}$$

Since consumer utility is linear in price and quality, there exists a well defined inverse

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

Moreover, $\dfrac{\partial \delta^{-1}(\delta_{ct}, q_{ct}, \xi_{ct})}{\partial q_{ct}}$ is constant. Note also that conditional on δ_{ct}, if there are no adjustment costs for (p_{ct}, q_{ct}), particular choices of (p_{ct}, q_{ct}) do not have dynamic implications. Therefore, producer problem can be written as

$$W(s_{ct-1}, s_{st-1}, x_{ct}) =$$
$$\max_{\delta_{ct}, q_{ct}} \left\{ s(s_{ct-1}, s_{st-1}, x_{ct}, \delta_{ct})(\delta^{-1}(\delta_{ct}, q_{ct}, \xi_{ct}) - C(q_{ct}, Z_{ct}) + \beta E[W(s_{ct}, s_{st}, x_{ct+1})] \right\}$$

From the first order conditions for optimal quality choice (given δ_{ct}),

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

where the first term is constant due to the linearity of the utility function it is clear that $q^*_{ct} = q(\xi_{ct}, X_{ct})$ does not depend on the optimal choice of a dynamic control, δ_{ct}.

The main implication of the discussion above is that if the vector of cost shifters Z_{ct} does not change across counterfactual simulations the optimal choice of quality will not change as well. Hence, changes in the market structure due to elimination of switching costs would affect only cable price. Note that this is an artifact of both the representative consumer model assumption and the linearity of the consumer per-period utility. For example, the argument would not work for a random coefficients demand side model.

Since according to the preliminary demand estimates the data does not reject a representative consumer model I simulate counterfactual scenarios using this model.

In order to calculate new cable prices I numerically solve static first order conditions for cable company where cost function is evaluated at observed vector of cost shifters and

[19] The same argument goes through in case if consumers are boundedly rational and forecast future flow utility using its current values only.

quality choices. The results of the counterfactual simulations suggest that in a static duopoly scenario with the unchanged satellite policy and the same vector of cable cost shifters average cable prices would be \$24.59, which is 28 percent higher than the observed average price. In case of a static monopoly scenario average cable price is \$28.90, which is about 51 percent higher than the observed value of \$19.20.

10 Conclusions

This paper addresses questions about the effects of consumer switching costs on the market structure and social welfare. Answers to these questions depend critically on the ability to estimate supply side parameters in presence of state dependent utility on the demand side. Multiple consumer types and/or products result in a state space that is too large to make numerical solution to the producer dynamic programming problem computationally feasible.

To overcome this problem I suggest to estimate supply side parameters from the optimality conditions for dynamic controls. In case when one dynamic control shifts several state variables at the same time it is very hard to derive an Euler equation in terms of the model primitives. However, there is a computationally feasible alternative that relies on simulation of the derivatives of the value function with respect to the state variables. Under the rational expectations assumption dynamic first order conditions could be used to form moment conditions for estimation.

The problem of unobserved distribution of consumer types across market shares and unobserved service characteristics is solved by first estimating the demand relationship and recovering the desired variables.

Preliminary estimates of the demand side parameters suggest switching costs of \$149 and \$238 for cable and satellite providers respectively. The demand side estimates cannot reject a representative consumer model, although mean levels of switching costs as well as price and quality coefficients are estimated precisely. Supply-side estimates imply the average cost of providing service per subscriber of about \$2.19 and average price-cost margin of \$16.52 per month.

Counterfactual simulations suggest that without switching costs in a duopoly situation average cable prices would be about \$24.59, which is 28 percent higher than the observed average prices of \$19.20. In case if there are no switching costs on the demand side and there are no satellite competitor (i.e. static monopoly scenario) average cable prices are estimated to be about \$28.90, which is 51 percent percent higher than the observed ones.

References

Bajari, P., Benkard, L., and Levin, J. (2007). Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370.

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

Berry, S. and Pakes, A. (2001). Estimation from the optimality conditions for dynamic controls. Unpublished manuscript. Yale University.

Borenstein, S. (1992). The evolution of u.s. airline competition. *The Journal of Economic Perspectives*, 6:45–73.

Che, H., Sudhir, K., and Seetharaman, P. (2006). Bounded rationality in pricing under state dependent demand: do firms look ahead? how far ahead? Unpublished manuscript. Yale School of Management.

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

Dube, J.-P., Hitsch, G., Rossi, P., and Vitorino, M. (2008). Category pricing with state-dependent utility. *Marketing Science*, 27(3):417–429.

FCC (2000). *Fact sheet. Cable television infromation bulletin.* Federal Communications Commission, 445 12th Street, S.W., Washington, D.C. `http://www.fcc.gov/mb/facts/csgen.html` (accessed March 01, 2009).

Gowrisankaran, G. and Rysman, M. (2007). Dynamics of consumer demand for new durable goods. Unpublished manuscript, University of Arizona.

Hansen, L. and Singleton, K. (1982). Generalized instrumental variables estimation of non-linear rational expectations models. *Econometrica*, 50(5):1269–1286.

Ho, C.-Y. (2009). Switching cost and the deposit demand in China. Unpublished manuscript. Georgia Institute of Technology.

Hotz, J., Miller, R., Sanders, S., and Smith, J. (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, 61(2):265–289.

Israel, M. (2003). Consumer learning about established firms: evidence from automobile insurance. Working Paper 0020, The Center for the Study of Industrial Organization at Northwestern University.

Kim, M., Kliger, D., and Vale, B. (2003). Estimating switching costs: the case of banking. *Journal of Financial Intermediation*, 12:25–56.

Kiser, E. (2002). Predicting household switching behavior and switching costs at depository institutions. *Review of Industrial Organization*, 20:349–365.

Knittel, C. (1997). Interstate long distance rates: search costs, switching costs and market power. *Review of Industrial Organization*, 12:519–536.

Krusell, P. and Smith, A. (1998). Income and wealth heterogeneity in the macroeconomy. *The Journal of Political Economy*, 106(5):867–896.

Melnikov, O. (2001). Demand for differentiated durable products: the case of the U.S. computer printer market. Unpublished manuscript, Cornell University.

Salies, E. (2005). A measure of switching costs in the GB electricity retail market. Working Paper 2005-9, GRJM.

Schiraldi, P. (2006). Automobile replacement: a dynamic structural approach. Job Market Paper, Boston University.

Sharpe, S. (1997). The effect of consumer switching costs on prices: a theory and its application to the bank deposit market. *Review of Industrial Organization*, 12:79–94.

Shcherbakov, O. (2008). Measuring consumer switching costs in the television industry. Unpublished manuscript, University of Arizona.

Shy, O. (2002). A quick-and-easy method for estimating switching costs. *International Journal of Industrial Organization*, 20:71–87.

Sturluson, J. (2002). *Consumer search and switching costs in electricity retailing.* Ph.D. thesis, Chapter 1, in Topics in the Industrial Organization of Electricity Markets, Stockholm School of Economics, Sweden.

Appendix 1. Estimates of the policy functions

Table 6: Linear regressions of cable price and quality on the set of state variables, OLS

Variable	Cab price	Cab quality
p_{st}	0.126	-0.035
	(0.028)	(0.008)
q_{st}	-0.077	0.035
	(0.036)	(0.011)
$MSO(p_{ct})$	0.489	0.005
	(0.011)	(0.003)
$MSO(q_{ct})$	0.184	0.876
	(0.056)	(0.017)
CAP_{ct}	-0.013	0.024
	(0.004)	(0.001)
$MCOAX_{ct}$	-0.277	0.591
	(0.152)	(0.046)
$MSO(CAP_{ct})$	-0.010	-0.021
	(0.007)	(0.002)
$MSO(SUB_{ct})$	-0.0007	-0.0086
	(0.005)	(0.002)
s_{ct-1}	-20.167	0.115
	(0.437)	(0.131)
s_{st-1}	3.963	-1.609
	(1.821)	(0.544)
ξ_{st}	-0.103	0.00003
	(0.052)	(0.015)
ξ_{ct}	5.493	-0.106
	(0.084)	(0.025)
$const.$	5.159	1.246
	(1.079)	(0.323)
R^2	0.73	0.66