Bundling and Vertical Relationships in Multichannel Television*

(Job Market Paper)

Ali Yurukoglu
Economics Department
NYU Stern

November 12, 2008

Abstract

I examine how banning the bundling of television channels would affect social welfare. The central innovation in this paper is accounting for the renegotiation of supply contracts between content providers, such as ESPN and CNN, and distributors, such as Comcast and DirecTV. I find that banning the bundling of television channels for sale to consumers causes wholesale input costs to rise. While consumers continue to receive the channels they value highly and pay a lower monthly bill, higher input costs foreclose consumers' access to some channels of moderate value. To generate these predictions, I use an industry model of viewership, demand, downstream pricing and bundling, and input market bargaining. I estimate the model's parameters using data on purchases, prices, ratings, and aggregate input costs. Next, I solve the model when downstream firms may not bundle channels, but instead must set prices for sizes of bundles and offer a few channels *à la carte*. Consumers may pick the exact channels they receive. Banning bundling changes the marginal surpluses generated by pairs of channels and distributors so that equilibrium input costs rise. I predict total welfare decreases by 2.02%, mean consumer surplus decreases by 0.96%, and industry profits decrease by 3.03%. By contrast, assuming input costs do not re-equilibrate to the policy change would predict mean consumer surplus to increase by 6.05% and total industry profits to decrease by 6.86%.

*This project benefitted enormously from the guidance of my dissertation advisors Ariel Pakes, Luis Cabral, John Asker, and Allan Collard-Wexler. I thank Heski Bar-Isaac, Adam Brandenburger, Gregory S. Crawford, Ignacio Esponda, Matt Grennan, Michael Katz, Robin Lee, Alessandro Lizzeri, Hong Luo, David Pearce, Vasiliki Skreta, Tracy Waldon, and the Stern micro workshop. Derek Baine of Kagan Research provided institutional detail on the multichannel television industry. I acknowledge the funding from the NYU Stern EMT department doctoral dissertation grant. I thank Norman White and the Stern Center for Research Computing for their help with the distributed computing infrastructure.**

I. Introduction

Distributors of multichannel television service offer consumers a short menu of bundles of television channels. Critics claim that bundling has raised prices, reduced choice, and therefore lowered welfare for the 115 million television households in the U.S. This paper uses an industry model and data to evaluate anti-bundling regulations considered recently by Congress and the Federal Communications Commission (FCC)[1]. The contribution here is to analyze the impact of anti-bundling regulations on the input market where distributors, such as Comcast and Time Warner Cable, contract with content providers, such as CNN and ESPN. I predict that consumer surplus, industry profits, and total welfare all decrease by between 1% and 3% if bundling is banned. The 2% decrease in total welfare corresponds to about 1.8 billion dollars (year 2000 dollars) per year. Without accounting for the re-equilibration of the input market, the model would have predicted consumer surplus to increase by roughly 6% (about 2.8 billion dollars per year), industry profits to decrease by 7% (about 3.3 billion dollars per year), and total welfare to slightly decrease.

Anti-bundling regulations would dramatically alter the choice set faced by the 115 million U.S. television households, 90% of whom subscribe to a multichannel television service. Therefore, most of society would benefit from careful analysis of these regulations. Still, there is no consensus on the likely effects of banning bundling. The lack of consensus is partly because anti-bundling regulations have not been implemented in enough similar circumstances to provide direct evidence. They have not been applied in this industry[2]. Experimentation is not practical due to the costs associated with experimenting on the necessary scale to evaluate the industry wide equilibrium effects[3]. With these options not available, this paper evaluates proposed policy changes using a model as a laboratory for the policy changes. This approach suffers from model error to the extent that the model does not capture reality. This paper's main contribution is to improve the model's accuracy by allowing the input market bargaining between downstream distributors and upstream channels to re-equilibrate to changing regulations [4].

This paper's key result is that equilibrium input costs, the fees paid by distributors to channels for

[1] *The Family and Consumer Choice Act of 2007* is a Congressional bill that would mandate that distributors allow consumers to pay a lower bill by opting out of channels they do not wish to receive. The FCC has studied anti-bundling regulations in FCC (2004) and FCC (2006). Furthermore, a group of consumers recently filed a class-action anti-trust case (Rob Brantley et al v. NBC Universal, Inc. et al) against cable and satellite companies seeking treble damages due to bundling and a requirement that channels are offered *à la carte*.

[2] Internationally, Canada and Hong Kong have introduced some anti-bundling regulations in multichannel television.

[3] Some local experimentation would be useful to gather evidence on how distributors would set prices to consumers.

[4] The multichannel television input market has been studied by Chipty (2001) using reduced form evidence, by Chipty and Snyder (1999) in the context of geographically separate downstream monopolies, and by Rennhoff and Serfes (2008) using degenerate distributions of consumer preferences and take-it-or-leave-it offers.

each subscriber that receives the channel, rise when renegotiated after bundling is banned. As input costs are the marginal costs for distributors, the increase is partially passed on as higher prices to consumers. The intuition for why input costs rise comes from distinguishing between consumers who have high valuations for individual channels and consumers who have high valuations for the bundle. Consumers with high valuations for the bundle may have low valuations for some individual channels included in the bundle. Each channel's equilibrium input cost roughly is proportional to the average valuation per subscriber to that channel. Average valuations for channels are higher when channels are not bundled together, because firms face steeper demand curves for each channel and set high prices.

I use an industry model grounded in institutional detail and historical data. The model builds on top of the demand, viewership, and pricing model in Crawford and Yurukoglu (2008) by modeling the choice by distributors of which channels to bundle together, and by modeling input costs as the outcome of a multilateral bargaining procedure. Here, I relax that paper's assumption that input costs do not change following anti-bundling regulations. The goal is thus to predict how input costs would respond when bundling is banned, and, ultimately, the effect on social welfare.

The model has three types of agents: consumers, downstream distributors, and upstream channels. Consumer behavior is determined by their preferences. Like Crawford and Yurukoglu (2008), I estimate those preferences using aggregate data on purchases, i.e. data on which bundle of channels consumers purchase and at what price, and data on viewership, i.e. which channels consumers watch and for how long. The viewership data provides a rich source of variation for estimating a flexible multivariate distribution of preferences for television channels. The purchase data provides empirical evidence about how households value channels relative to income.

On the supply side, downstream firms compete with each other and negotiate input costs with upstream firms. I assume that downstream firms compete by choosing prices and bundles. I assume that observed prices and bundles are a Nash equilibrium given estimated preferences. I estimate input costs as those which make the Nash equilibrium assumption hold. I use the procedure in Pakes, Porter, Ho, and Ishii (2006) to incorporate a subset of necessary conditions of Nash equilibrium in bundle choice into the estimation. This restricts estimated input costs to reflect that adding or dropping a channel from an observed bundle should reduce profits on average for the firms making the decision. Modeling pricing and bundling jointly places this paper within a recent empirical literature studying firms' choices of non-price attributes, often called "product positioning," such as Seim (2006), who studies location decisions for video rental stores, and Crawford and Shum (2006), who study the choice by cable television firms of how many channels to offer, but, unlike this paper, not the identity of those channels.

To model the determination of input costs, I fix an industry bargaining protocol similar to the models of Horn and Wolinsky (1988) and de Fontenay and Gans (2007). The bargaining protocol features bilateral meetings between channels and distributors whose outcomes impose externalities on other firms due to downstream competition. One notable empirical paper that also studies bargaining with externalities due to downstream competition is Ho (2008) who studies hospital-HMO negotiations in the U.S. This paper contributes to this line of research by using a general bargaining protocol that includes Ho's take-it-or-leave-it offers as a special case. I estimate pair-specific bargaining parameters that produce the estimated input costs in equilibrium.

Having estimated preferences and bargaining parameters, I am in a position to re-solve the model when downstream firms may no longer bundle. However, I first examine the estimated parameters to see how variation in the data comes through with the extra structure of the model imposed and how the estimates compare with previous estimates of similar parameters.

Key features of the ratings data manifest in the estimated distributions of household preferences for channels. Highly watched channels such as ESPN and TNT have higher estimated unconditional mean willingness to pay (WTP) than less watched channels. Black households are estimated to have a higher mean WTP for Black Entertainment Television (BET) than non-black households. Older households have higher estimated mean WTP for 24 hour news and infotainment channels, such as CNN and Fox News. Family households likewise have higher estimated mean WTP for the children's networks Nickelodeon and Disney Channel. The estimated mean own-price elasticities for television service are -1.86 for basic service, -5.276 for expanded basic service, and -4.33 for satellite service. I find that expanded basic cable, digital basic cable, and satellite service are more substitutable with each other than with basic cable or no multichannel television service.

I estimate that large distributors such as Comcast face about %10 lower input costs than small and independent operators. I estimate that vertical integration between channels and distributor does not affect input costs for the integrated distributor relative to other distributors. This is partly because I focus on established channels for my analysis. A cursory analysis of the raw data suggests that vertical integration is important for new or small channels, but this is not true for large and established channels. Neither prices nor carriage are systematically different for distributors who are vertically integrated with established channels.

The estimated bargaining parameters reject take-it-or-leave-it offers as a model of the input market for this industry. On average, distributors are estimated to have higher bargaining parameters than channels. Within distributors, estimated bargaining parameters are higher on average for big cable firms than for small cable firms and satellite. Even though small cable's estimated input costs are higher than satellite's, their estimated bargaining parameters are still higher than satellite's.

I simulate the medium-run welfare effects of forcing distributors to price sizes of bundles of channels, called Bundle Size Pricing (BSP), as in Chu, Leslie, and Sorensen (2006) and to offer a few famous channels *à la carte*. In BSP, distributors set prices for bundles of 5, 20, and 60 channels. Consumers may then choose a size, if any, and fill their bundle with the channels of their choosing. In addition, they may select to subscribe to any subset of the channels offered *à la carte*. I find that while consumer welfare would increase if input costs were held fixed, input costs go up which exacerbates the double marginalization problem. Higher input costs result in minor decreases for all three measures: total welfare, mean consumer surplus, and industry profits. Consumer welfare gains survive for a subset of consumers. On average, consumer surplus is substantially lower than if the input market were held fixed at the out-of-equilibrium original input costs.

II. The Multi-Channel Television Industry

The multichannel television market is a two-sided market. Cable and satellite distributors[5] provide a platform connecting households to content providers. When consumers watch programs, their consumption creates audiences. Channels sell audiences to advertisers.

Figure 1 provides a graphical representation of the supply chain by which programming is produced and sold to households, and audiences are created and sold to advertisers. Downward arrows represent the flow of programming from Content Providers to Households.[6] Upward arrows represent the creation and sale of audiences to advertisers. The various sub-markets that characterize the purchase and sale of content or audiences are indicated at each step in the chain. In this paper, I focus on the for-pay distribution market, the advertising market, and the programming market.

A. *The MVPD Market*

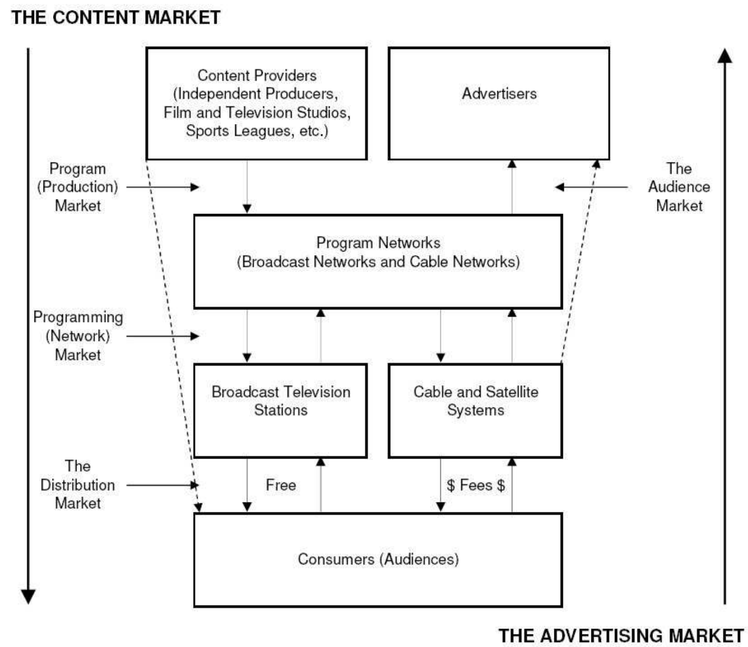
Multichannel Television Services: Bundles of Program Channels Cable television systems choose a portfolio of television channels, bundle them into services, and offer these services to consumers in local markets. Satellite television systems similarly choose and bundle channels into services, but offer them to consumers on a national basis. Almost all cable and satellite firms offer nonlinearly priced packages that combine multichannel television service with internet service and telephone service.

Cable and satellite systems offer four types of channels. *Broadcast networks* are advertising-supported television signals broadcast over the air in the local market by television stations and

[5] In regulatory proceedings, these firms are multichannel video program distributors (MVPD).

[6] The distribution rights to content (e.g. a television program like "Crocodile Hunter") is purchased by a Television Channel (e.g. CBS or The Discovery Channel) and allocated a time slot in its programming lineup. These channels are then distributed to consumers in one of three ways. Broadcast Networks, like ABC, CBS, and NBC, distribute their programming over the air via local broadcast television stations at no cost to households. Cable channels like The Discovery Channel, MTV, and ESPN distribute their programming via cable or satellite television distributors that charge fees to consumers. The dashed arrow between content providers and consumers represents the small but growing trend to distribute some content directly to consumer via the Internet.

Figure 1. Television Programming Industry



then collected and retransmitted by cable systems. Examples include the major, national broadcast networks – ABC, CBS, NBC, and FOX – as well as public and independent television stations. *Cable programming channels* are advertising- and fee-supported general and special-interest channels. Examples include some of the most recognizable brands in media and entertainment, such as MTV, ESPN, CNN, Nickelodeon, and The Discovery Channel. *Premium programming channels* are advertising-free entertainment channels such as HBO and Showtime. *Pay-Per-View* are devoted to on-demand viewing of high-value programming, typically the most recent theatrical releases and specialty sporting events.

Cable and satellite systems exhibit moderate differences in how they bundle channels into services. Broadcast networks and cable channels are typically bundled and offered as *Basic Service* while premium programming channels are typically unbundled and sold as *Premium Services*.[7] Beginning in the mid-1990's, systems further divided Basic service, offering some portion of their cable channels

[7]In the last 5 years, premium channels have begun "multiplexing" their programming, i.e. offering multiple channels under a single brand (e.g. HBO, HBO 2, HBO Family, etc.).

on multiple services, called *Expanded Basic* and *Digital Services*. For either Basic or Expanded Basic Services, consumers are not able to buy access to the individual channels offered in bundles; they must instead purchase the entire bundle.

Regulation in Multi-Channel Television Markets Multichannel television distributors are subject to a number of regulations impacting channel carriage and bundling decisions, prices, and other features of these markets.

The specific content of any cable service may not be regulated on First Amendment grounds. However, the 1992 Cable Act introduced two regulations that impact the channels offered on a cable system and how they are bundled into services. First, the Act required the creation of a Basic tier of service containing all offered broadcast and public-interest programming carried by the system. This Basic Service may also include some or many cable programming channels, at the discretion of the system. Many systems responded by introducing bare-bones "Limited Basic" services containing only those channels they were required to offer. Second, the Act introduced Must-Carry/Retransmission Consent. These regulations give local broadcast stations the option either to demand carriage on local cable systems (Must-Carry) or negotiate with those systems for compensation for carriage (Retransmission Consent).[8]

The 1992 Cable Act re-introduced price regulation into cable television markets. Regulation differed by tiers of cable service and only applied if a system was not subject to "effective competition."[9] Basic tiers were regulated by the local authority, which was required to certify with the FCC. Higher tiers were regulated by the FCC. Regulation of higher tiers, however, was phased out by the 1996 Telecommunications Act and eliminated as of March 31, 1999. Regulation of Basic Service rates in areas of little competition remains the only source of price regulation in the cable industry.

In the programming input market, cable and satellite distributors negotiate carriage agreements for channels on a bilateral basis between a cable channel, or a group of cable channels, and a distributor. Distributors with several local systems are commonly known as Multiple System Operators (MSOs). These agreements specify transfers between the two parties and terms of carriage such as which tier the channel will be on. The 1992 Cable Act introduced rules that forbid vertically integrated cable and satellite systems and channels from discriminating against unaffiliated rivals in either the programming or distribution markets. Carriage agreements have "Most Favored Nation" clauses that standardize terms between channels and cable systems of a given size.

There have been fewer regulations in the satellite television market. The Satellite Home Viewer

[8]Smaller (esp. UHF) stations commonly select Must-Carry, but larger stations and station groups, particularly those affiliated with the major broadcast networks, have used Retransmission Consent to obtain compensation from cable systems, often in the form of carriage agreements for broadcaster-affiliated cable channels.

[9]See Crawford (2006) for a survey of the history of price regulation in cable television markets.

Improvement Act (SHVIA) was passed on November 28, 1999. It permitted satellite providers to distribute local broadcast signals within local television markets.[10] This leveled the playing field between cable and satellite systems and established the latter as an effective competitor in U.S. multi-channel television markets.[11] Since 2002, satellite systems that distribute local signals must follow a "carry-one, carry-all" approach similar to Must-Carry and must negotiate carriage agreements with local television stations under Retransmission Consent (FCC (2005)). Unlike cable systems, satellite providers have never been subject to price regulations.

B. *The Advertising Market*

Most advertising space is sold by channels, but three or four minutes per hour are sold by the local cable system.[12] Individual channels vary in what percentage of revenues come from advertising versus subscription fees. On the aggregate, advertising revenues account for nearly one half of total channel revenues. For particular channels, advertising revenues depend on the total number and demographics of viewers and subscribers. These figures, called ratings, are measured by Nielsen Media Research. Ratings are measured at the Designated Metropolitan Area (DMA) level, of which there are 210 in the United States. In urban areas, the DMA usually corresponds to the greater metropolitan area. DMA's usually include multiple cable systems, often from different owners. For local advertising purposes, these systems are allowed to join together to form an "interconnect" which allows advertisers to reach multiple local systems within a DMA.

III. The Data

This section describes the data underlying this study. I divide the data into two categories: market data, which measure consumers' purchasing decisions or firms' production decisions, and viewership data, also called ratings, which measure consumers' utilization of the cable channels available to them.

[10] Within a year, satellite providers were doing so in the top 50-60 television markets. They now do so in almost all television markets, allowing them to provide a set of services comparable to those offered by cable systems for the vast majority of U.S. households.

[11] Every net new subscriber to multi-channel television markets since 2000 has been a satellite subscriber. See Crawford (2006) for details. Satellite firms are taxed differently than cable firms.

[12] Kagan reports local advertising revenue to cable systems for 2006 of approximately \$3.7 billion, 5.1% of total cable system revenue.

A. *Market Data*

Market data in the MVPD industry comes from two sources: Warren Communications and Kagan Research. Warren produces the monthly Television and Cable Factbook Electronic Edition (henceforth Factbook). The Factbook provides data at the cable system level on prices, bundle composition, quantity, system ownership and other system characteristics. Kagan produces the Economics of Basic Cable Networks yearly (henceforth EBCN). EBCN provides data at the channel level on revenue, cost, and subscriber quantities.

Factbook and Satellite Data The Factbook sample spans the time period 1997-2007. The Factbook collects the data by telephone and mail survey of cable systems. The key data in Factbook are the cable system's bundle compositions, the prices of its bundles, the number of monthly subscribers per bundle, and ownership. Various subsamples of the Factbook have been used in previous studies of the MVPD industry.[13]

Table 1 provides summary statistics for the Factbook data. An observation is a system-bundle-year (e.g. NY0108's Expanded Basic in 2000). I observe data on over 20,000 system-year-bundles, based on almost 16,000 system-years from over 6,800 systems. Most systems in the data offer a single Basic bundle, while the majority of the rest offer just Basic + Expanded Basic service. While currently rare, much of the data comes from early in the sample period when fewer offerings were the norm.

For each of these bundles, and by market type, Table 1 reports the average price of the bundle (in year-2000 dollars), its market share, and the number of cable channels offered. Systems offering multiple services differentiate them with respect to total channels and price: while the average Basic service in our data costs $24.14 and offers 17.4 cable channels, the average Digital Basic bundle costs $48.33 and offers 81.2 channels.[14]

One important feature of the Factbook data is the variation in composition of bundles, both within and across markets. Summary statistics on channel carriage is available in Crawford and Yurukoglu (2008). Cable systems tailor their bundles to their market given their varied wholesale costs of channels. For example, ESPN is carried by almost all systems (97%) in our data. Of these, most (77%) carry it on Basic Service. By contrast, smaller channels are frequently offered on a Digital Service.

[13] To name only a few: Crawford (2000), Chipty (2001), and Crawford and Yurukoglu (2008). The last paper presents further detail on the Factbook data set.

[14] Digital basic packages were made possible by cable systems investments in digital infrastructure in the late 1990's and 2000's. This increased the bandwidth available for delivering television channels. Prior to digital upgrades, most systems offered simply a basic bundle or a basic bundle and an expanded basic bundle. Following the digital upgrades, many systems also offered a higher tier, called digital basic, and, sometimes, a digital expanded basic bundle, as well as Internet services and high definition channels.

I also use market data on satellite television service. Unlike for cable service, these do not vary by geography.[15] This information was collected by hand.[16] I then matched this to aggregate satellite penetration data, $\frac{totalsatellitesubscribers}{totaltvhouseholds}$, at the DMA level from Nielsen Media Research. Table 2 provides price and total channels information by year for the DirecTV Total Choice package.

Kagan (ECBN) Data I use the 2006 edition of the EBCN. The sample covers 120 cable channels with yearly observations dating back to 1994 when applicable. The key variables are total subscribers, license fee revenue, advertising revenue, and ownership. The key variable in this study is average license fee per subscriber. For example, ESPN may have negotiated one fee with Comcast and another fee with Cox, a smaller distributor. The EBCN average is, for each channel, over distributors who may pay different license fees. The EBCN data are collected by survey, private communication, consulting information, and some estimation. The exact methods used are not disclosed. Summary statistics for those data are provided in Table 3. EBCN has been used in fewer MVPD industry studies than Factbook.[17]

B. *Viewership Data*

The viewership data comes from Nielsen Media Research. Television ratings data is collected by different methods depending on the market and type of data. I use tuning data from the 56 largest DMA's for about 65 of the biggest cable channels over the period 2000-2006 in each of the months of February, May, July, and November (known for historical reasons as the sweeps months). The main variables are the DMA, the program, the channel, the program's rating, and the channel's cumulative rating. The rating is the percentage of television households in the DMA viewing the program. The channel's cumulative ratings ("cume") indicates what percentage of television households with access to the channel tuned to the channel for at least ten minutes in a given week. Nielsen data is used throughout the television industry for scheduling of programs and pricing of advertising. Previous academic studies using similar data include Hausman and Leonard (1997) and Crawford and Yurukoglu (2008).

I aggregate the information across programs on each channel within each month. Thus an observation is a channel-DMA-year-month. There are 1,482 such combinations. Table 4 presents summary statistics for a subset of channels. It demonstrates that there is considerable variance in the monthly DMA average ratings both within and across channels. Ratings are highest for the most widely available channels, although this pattern is not monotonic. For example, The Hallmark Channel is the 41st most widely available channel, but has the 27th highest rating). Highly rated channels

[15] Save for the carriage of local broadcast signals.
[16] I also compared with a data set used by (Chu 2006) to reduce measurement error.
[17] Chu (2006) and Kagan's own commercial research.

typically have higher average cumes.

Channels' ratings vary from DMA to DMA and across time. Two important types of across-DMA and time variation I use are (1) how ratings vary with the demographic composition of a DMA and (2) how ratings co-vary (conditional on demographic differences). I use eight demographic characteristics: Urban/Rural status, Family status, Income, Race (White/Black/Hispanic/Asian), Education, and Age. As an illustrative example of the impact demographic characteristics can have on ratings, I present a graph of the ratings of Black Entertainment Television (BET) in its least popular and most popular DMA's for 2004 in Figure 2. Unsurprisingly given the target audience of BET, the channel has its highest ratings in heavily black populated DMA's such as Memphis and its lowest ratings in sparsely black populated DMA's such as Salt Lake City. The share of black population is an important predictor of ratings for BET. This paper's estimates of willingness-to-pay for BET are higher for black households directly as a consequence of this pattern in the data.

C. *Data Quality*

In the data quality appendix to Crawford and Yurukoglu (2008), we call attention to the nonstandard features of these data sets, particularly the Factbook. There, we focus on missing market share and price data. About two thirds of the possible observations on market share and price for cable bundles are either missing, not updated from the previous year, or both. For this paper, I maintain the assumption that this data is missing at random conditional on the observable characteristics of the system.

IV. Multichannel Television Industry Model

The industry model predicts demand for multichannel television services, household viewership of channels, prices and bundles offered by distributors, and distributor-channel specific input costs. This section derives those predictions in terms of a variable set of parameters. The next section, on identification, estimation, and inference, picks a particular set of parameters so that the predictions from the model align with their empirical counterparts.

The timing of the model is:

Stage 1 Channels and distributors bargain bilaterally to decide input costs.

Stage 2 Distributors set prices and bundles.

Stage 3 Households make purchases.

Stage 4 Households view television channels.

I start from the last stage and work backwards.

A. *Demand and Utilization of Multichannel Television Services*

I model households choosing what bundle of channels, if any, to subscribe to, and how much time to allocate for watching different channels conditional on choice of bundle. This part of the model is the same as in Crawford and Yurukoglu (2008).

Household i in market m faces a set of bundles $\mathbf{b_m}$ offered by the firms operating in market m. A bundle $j \in \mathbf{b_m}$ consists of a set of channels C_j, a list of other attributes z_j which include the name of the firm offering bundle j, the price of the bundle p_j, and a scalar ξ_j which measures the utility generated by attributes of bundle j which are known to the household and firms, but unobserved to the econometrician. A typical market would have two or three bundles from a cable firm, and three to five bundles from each of the satellite firms[18]. I denote the set of bundles offered by firm f in market m as $\mathbf{b_{fm}}$.

Conditional on subscribing to bundle j, household i allocates its leisure time to watching the channels in C_j, including the figurative channel 0 which represents non-television leisure, according to the following optimization problem:

$$\begin{aligned} \max_{\mathbf{t_{ij}}} \quad & \textstyle\sum_{c \in C_j} \gamma_{ic} \log(t_{ijc}) \\ \text{subject to} \quad & \textstyle\sum_{c \in C_j} t_{ijc} \leq T_i \end{aligned} \tag{1}$$

where γ_{ic} is household i's preference parameter for watching channel c, t_{ijc} is the amount of time household i tunes into channel c when subscribing to bundle j, and T_i is household i's total leisure time.

The assumed functional form says the more a household values a channel, the more it watches the channel. Similarly, if a household never watches a channel, it values that channel at zero. As an example for when this implication would be undesirable, consider a hurricane prone region and The Weather Channel. If no hurricanes are threatening the region, very few households would watch The Weather Channel. However, they might value the option of having The Weather Channel in case a hurricane develops nearby[19].

[18] In this paper, due to data limitations, I assume only one satellite offering when I estimate demand.

[19] In the ratings data, viewership of The Weather Channel soared during Hurricane Katrina. Since I aggregate over time in estimating preferences, I will partially capture some option value.

The indirect utility, v^*_{ij}, given by evaluating the objective function at optimal channel times t^*_{ijc} is:

$$v^*_{ij}(\gamma_{\mathbf{i}}, T_i, C_j) = \sum_{c \in C_j} \gamma_{ic} \log(\frac{\gamma_{ic}}{\sum_{c \in C_j} \gamma_{ic}} T_i) \tag{2}$$

Estimation will partly be based on matching the models predicted conditional variances of time watched for channels with observed conditional variances of ratings.

Utility from viewing channels is only a part of the household's utility from subscribing to the bundle. For example, a household suffers from the disutility of paying for the bundle. I assume the total monthly utility household i enjoys from subscribing to bundle j, suppressing market and month subscripts, is:

$$u_{ij} = v^*_{ij} + z'_j \lambda + \alpha_i p_j + \xi_j + \sigma_\epsilon \epsilon_{ij} \tag{3}$$

where λ and α_i are taste parameters for non-channel attributes of bundle j and price, respectively. ϵ_{ij} is a household-bundle specific utility shock which is assumed to be distributed extreme value with mean zero and variance one, so that $\sigma_\epsilon \epsilon_{ij}$ is distributed extreme value with mean zero and variance σ^2_ϵ.

I assume the presence of an outside good whose utility is normalized to zero. The interpretation of the utility household i enjoys for bundle j is therefore the difference in utility enjoyed by household i between bundle j and not subscribing to any bundle. Given the distribution of parameters of the utility function and the attributes of the bundles in a given market, I can compute the model's predicted market shares by aggregating over utility maximizing households.

I first introduce some commonly used notation which will make the expressions more compact.

$$\delta_j = z_j \lambda + \alpha p_j + \xi_j$$

where $\alpha = \int \alpha_i$, and

$$\mu_{ij} = v^*_{ij} + (\alpha_i - \alpha) p_j$$

so that

$$u_{ij} = \delta_j + \mu_{ij} + \sigma_\epsilon \epsilon_{ij}$$

Let A_j be the set of households for whom bundle j delivers the highest utility from the set of bundles available, including the outside good, without the idiosyncratic taste shock ϵ_{ij}.

$$A_j = \{i | \delta_j + \mu_{ij} \geq \delta_k + \mu_{ik} \forall k \in B_m\} \quad (4)$$

The model's predicted market share for bundle j in market m in month t, when F^m is the distribution function of households in m, is given by

$$s_{jmt} = \int_{A_{jmt}} \frac{(exp(\delta_{jmt} + v^*_{ijmt})\sigma_\epsilon^{-1}) dF^m(i)}{\sum_{k=0}^{|B_{mt}|} exp((\delta_{kmt} + v^*_{ikmt})\sigma_\epsilon^{-1})} \quad (5)$$

Estimation will be based partly on setting these predicted market shares equal to their empirical counterparts.

B. *Supply: Downstream Distributors*

Distributors compete by choosing the composition and price of their bundles to maximize profits. I assume that observed prices and bundles form a Nash equilibrium of the price and bundle choice game.

The profit of a distributor before fixed costs is:

$$\Pi_{fm}(\mathbf{b_m}, \mathbf{p_m}) = \sum_{j \in \mathbf{b_{fm}}} (p_j - \sum_{c \in C_j} \tau_{fc}) D_{jm}(\mathbf{b_m}, \mathbf{p_m}) \quad (6)$$

where f denotes firm, m market, and j bundle. $\mathbf{b_m}$ is a list of offered bundles in market m with corresponding prices $\mathbf{p_m}$. τ_{fc} are firm-channel specific carriage fees. Firm f pays channel c a fee of τ_{fc} for every household which receives channel c from firm f. The set of bundles offered by firm f is $\mathbf{b_{fm}}$. The set of channels in bundle j is C_j.

Separate the bundles offered in market m into those offered by firm f and not: $\mathbf{b_m} = (\mathbf{b_{fm}}, \mathbf{b_{-fm}})$. The same for prices: $\mathbf{p_m} = (\mathbf{p_{fm}}, \mathbf{p_{-fm}})$. Nash equilibrium assumes:

Nash Assumption $\forall f$ and $\forall m, \mathbf{b_{fm}}$ and $\mathbf{p_{fm}}$ maximize $\Pi_{fm}(\mathbf{b_m}, \mathbf{p_m})$ given $\mathbf{b_{-fm}}$ and $\mathbf{p_{-fm}}$.

The Nash assumption implies that bundle prices satisfy the firm's first order necessary conditions for maximizing profit. Furthermore, if an observed bundle is modified by adding or removing a channel, then the profit will be less than or equal to the original bundle's profit, no matter the price of the new bundle. Identification and estimation of input costs is partly based on these implications of the Nash assumption.

I do not have a uniqueness result for the Nash equilibria of this pricing and bundling game. The estimation of input costs relies only on the necessary conditions of Nash equilibrium. Therefore, multiple equilibria does not affect the properties of the estimated cost parameters. Multiple Nash equilibria would hinder both the estimation of bargaining parameters and the simulation analysis of unrealized policies. While I can not prove uniqueness, I do numerically search for multiple equilibria by changing initial prices. I always fail to find more than one.

C. *Supply: Channel Distributor Negotiations*

Input costs are the outcome of bilateral negotiations between upstream channels and downstream distributors. Bilateral negotiations have been studied extensively building on Nash (1950) and Rubinstein (1982), as detailed in Muthoo (1999). Chipty and Snyder (1999) use such models to analyze mergers in the multichannel television industry before the emergence of satellite television or cable overbuilds. This paper's environment differs from those models because payoffs depend on outcomes of bilateral negotiations that firms are not party to. These externalities are due to downstream competition. Horn and Wolinsky (1988), Hart and Tirole (1990), McAfee and Schwartz (1994), and Segal and Whinston (2003) study these environments when one side of the market has one or two agents. Raskovich (2003) extends these models to capture the notion of pivotal buyers in the multichannel television industry. Prat and Rustichini (2003) and de Fontenay and Gans (2007) extend these models to allow for arbitrary numbers of agents on both sides of the market.

I too model this situation as a game involving the upstream channels and the downstream distributors. Distributors and channels meet bilaterally. They bargain according to a fixed protocol to determine whether to form an agreement, and if so, at what input cost. The ultimate payoffs are determined by downstream competition at the agreed upon input costs.

I assume that the agreements between channel and distributor are simple linear fees: how much must the distributor pay to the channel each month for each subscriber who receives the channel. In reality, the contracts are longer. Many contain marketing agreements, most favored nation clauses, and tier placement requirements. Few contain non-linear monetary transfers for distributors of a fixed size. I model the contracts as only a linear fee for each pair[20].

In the bargaining stage, each pair of channel and distributor meets separately and simultaneously. I assume these meetings result in the asymmetric Nash bargaining solution[21]. Let $\Psi = \{\tau_{fc}\}$ be a set of input costs, a scalar for each pair of distributor and channel. If there is no agreement between a distributor and a channel, then the input cost is positive infinity. In each bilateral meeting, τ_{fc} maximizes firm f and channel c's bilateral Nash product:

$$NP_{fc}(\tau_{fc};\Psi_{-fc}) \quad = \quad \Big[\Pi_f(\tau_{fc};\Psi_{-fc}) - \Pi_f(\infty;\Psi_{-fc})\Big]^{\zeta_{fc}}\Big[\Pi_c(\tau_{fc};\Psi_{-fc}) - \Pi_c(\infty;\Psi_{-fc})\Big]^{1-\zeta_{fc}}$$

Negotiations are simultaneous and separate, so Ψ_{-fc}, the set of all other input costs, is not known but conjectured. ζ_{fc} is the bargaining parameter of distributor f when meeting channel c. Allowing $\zeta_{fc} \neq 0.5$ distinguishes asymmetric Nash bargaining from symmetric.

Bargaining Equilibrium $\forall f$, c, τ_{fc} maximizes $NP_{fc}(\tau_{fc};\Psi_{-fc})$ given Ψ_{-fc}.

The interpretation of this equilibrium is a Nash equilibrium between Nash bargains as in Horn and Wolinsky (1988). Consider a simultaneous move game where the players are the bargaining pairs and their payoffs are their Nash products. The bargaining equilibrium is the Nash equilibrium of that game.

One issue, also raised in Horn and Wolinsky (1988) and discussed in Raskovich (2003), is how to define the disagreement payoffs. Following the Nash equilibrium reasoning, I assume that agreements are binding in all contingencies. In the appendix, I solve an alternative case where if a pair disagrees, all other firms re-negotiate conditional on the disagreeing pair dropping out forever. This case is reminiscent of the reasoning in the Shapley value[22]. This paper's conclusions do not depend on which assumption I choose.

I will now use the connection between Nash bargaining and Rubinstein bargaining from Binmore (1987) to draw attention to the bargaining model's assumptions from a non-cooperative point of view. Suppose that instead of Nash bargaining, each pair meets and plays a repeated game. In each period, nature chooses one of the two sides to propose a τ_{fc} to the other. If the non-proposing

[20] Linear input costs above the production marginal cost, in this case zero, are often considered unrealistic because with downstream monopoly, the upstream and downstream firms can find fixed transfers that make both better off after changing the input cost to marginal cost. However, when there is downstream competition, committing to linear contracts is one way of avoiding the dissipation of profits due to downstream competition.

[21] I will use the connection between the asymmetric Nash bargaining solution and the limiting stationary equilibrium of random proposer Rubinstein bargaining game.

[22] de Fontenay and Gans (2007) make an explicit connection with a cooperative solution that has the flavor of the Shapley value.

player accepts, then the negotiations end. Otherwise, nature either ends the pair's negotiation with a small probability, or the game continues. Firms are engaged in several negotiations at once, each of which affects every player's ultimate payoff. Therefore, this is a game of asymmetric information. The bargaining equilibrium I consider above corresponds to the limit of stationary subgame perfect equilibria of this game as the probability of exogenous breakdown approaches zero, and players hold passive beliefs when evaluating out-of-equilibrium offers. Passive beliefs assumes that when a player receives an out-of-equilibrium offer in a negotiation, it believes that its bargaining partner is continuing its negotiations with their other partners at the anticipated equilibrium levels.

I treat each channel as an individual firm even though channels are often part of larger conglomerates. For example, ABC Family, The Disney Channel, ESPN, ESPN2, ESPN Classic, SOAPNet, and several other channels are currently under the ownership of The Walt Disney Company. I assume that that the disagreement points for each of these channels are the profits from only that channel being dropped, rather than from all or a subset of channels from the conglomerate being dropped.

Passive beliefs are attractive due to their simplicity to implement. They are unattractive because they do not capture some features of negotiations that are appealing in certain situations[23]. One stark example is given by Segal and Whinston (2003). They consider an upstream firm with a commonly known capacity constraint who deviates by offering a particular downstream firm a quantity at the capacity constraint. The downstream firm should know that it is technologically impossible that its rivals are supplied by the upstream firm, and so passive beliefs require the downstream firm believe in the impossible. In the current setting, consider a pair, say ESPN and Comcast, meeting to negotiate an input cost. Passive beliefs can be interpreted as each firm sending a different representative to the meeting, and these representatives do not talk to the other representatives of its own firm. That is, even though Comcast's payoffs depends on its negotiated costs with other channels, passive beliefs require that Comcast's representative ignores this information when dealing with ESPN. By using passive beliefs to ignore asymmetric information[24], the model sacrifices capturing incentives due to informational asymmetries, but gains tractability in determining how the threat of unilateral disagreement determines input costs in an bilaterally oligopolistic setting.

[23] Furthermore, Rey and Verge (2004) demonstrate another feature of using passive beliefs that could be problematic. They show that an equilibrium supported by passive beliefs may not exist in settings where an equilibrium supported by other belief systems, such as 'wary' beliefs, does exist.

[24] As a separate issue, I also ignore moral hazard issues. For example, I ignore the imperfectly observable choice of effort exerted by channels into making compelling programming following an agreement.

D. *Limitations of Model*

This industry model is static. I do not model the dynamic decisions in the industry. These may be related to learning[25] or habit formation on the demand side, and investment, entry, or exit on the supply side. Rather, I assume that these decisions have been made and will not change.

V. Identification, Estimation, and Inference of Model

I separate estimation into three steps: demand estimation, cost estimation, and bargaining power estimation. It would be more efficient to estimate the whole model jointly. It is computationally and conceptually more simple to separate the estimation. I use the necessary conditions of Nash equilibrium price setting in both demand and cost estimation.

A. *Identification and Estimation of Demand*

I estimate the demand side of the model following in the tradition of Berry, Levinsohn, and Pakes (1995). Unlike most BLP-type demand models, I also make use of ratings data which measures how long households watch various channels. This extra source of information, combined with the assumptions of the household viewership model, helps estimate more plausible distributions of household preferences than one could hope for with only aggregate purchase data.

Heuristically, I, as in Crawford and Yurukoglu (2008), estimate preferences by rewriting v^*_{ij} as the dot product of a vector of household-bundle specific channel taste parameters, β_{ijc}, and a vector of dummy variables, x_{jc}, indicating whether channel c is in bundle j. I then parameterize β_{ijc} as a function of household i's demographics, i's unattributable taste parameters, and the channels available on bundle j. I estimate the distribution of β_{ijc} in the population straight off of variation in the ratings data. I then construct v^*_{ij} and choose λ and the distribution of α_i so that the predicted market shares of bundles in the model match the observed market shares in the data, accounting for the endogenous relationship between p_j and ξ_j using an instrumental variables assumption as in Berry (1994).

I start by writing

$$\begin{aligned} v^*_{ij} &= x_j \beta_{ij} \\ \beta_{ijc} &= \gamma_{ic} \log(t_{ijc}) \qquad c \in C_j \\ & \quad 0 \qquad c \notin C_j \end{aligned}$$

[25] In this industry, learning is commonly part of "channel surfing."

I now make a parametric assumption on the vector of $\beta_{ij} = (\beta_{ij1}, \beta_{ij2}, \ldots, \beta_{ij|C|})$:

$$\beta_{ij} \quad = \quad \beta + \Pi D_i + v_i + \Theta \widetilde{x_j} \tag{7}$$

The vector of channel taste parameters for household i in bundle j has a linearly separable functional form with terms that capture demographic influences, unattributable tastes (v_i, distributed in the population according to distribution function $G(v_i)$), and the influence of bundle composition ($\Theta\widetilde{x_j}$ where each row of $\widetilde{x_j}$ indicates if bundle j has specific channels.). For example, the row of $\widetilde{x_j}$ corresponding to Fox Soccer Channel would include a component that is a dummy variable for whether bundle j has GolTV, another 24-hour soccer channel. One would expect the coefficient in Θ for this term to be non-positive, indicating that these channels are substitutes. At the same time, one would expect the correlation between those channels' dimensions of v_i's to be positive, indicating that households who have high unattributable tastes for one also have high unattributable tastes for the other.

I observe aggregate ratings data which I associate with predicted aggregate viewership from the model. Households' selection into bundles is implicit in the aggregation of observed data; Nielsen samples randomly from the population, but observes viewership for the sampled households only at the bundle chosen by the household. The goal for the moment is to estimate the unconditional distribution of β_{ij} using the conditional observed distribution of ratings and demand covariates which affect selection, but without computing the demand model.

I apply the Nielsen sampling operator, Υ^{dm}[26], to both sides of 7. Υ^{dm} operates on a data set of all households in DMA d and month m by taking the Nielsen sample household average for DMA d and month m. For example, applying Υ^{dm} to the data set of all household incomes in DMA d and month m would produce the average income of households sampled by Nielsen in DMA d and month m. As discussed above, when operating on data which is a function of households' chosen bundles, the Nielsen operator produces a conditional sample average.

$$\Upsilon^{dm}\beta_{ij} \quad = \quad \beta + \Pi D_d + \Upsilon^{dm} v_i + \Theta \widetilde{x_m}$$

I then take the Nielsen sample average of the other side of the equation: $\gamma_{ic}\log(t_{ijc})$. After the approximations spelled out explicitly in Crawford and Yurukoglu (2008), because of the equality of the two aggregated terms, I have:

[26] $\Upsilon^{dm} = \frac{1}{N_{dm}} \sum_{i \in \text{Nielsen sample of DMA d and month m}}$ where N_{dm} is the number of households in the Nielsen sample of DMA d and month m. Υ^{dm} is a linear operator.

$$r_{dm}\log(r_{dm}T) \quad \approx \quad \beta + \Pi D_d + \Upsilon^{dm} v_i + \Theta\widetilde{x_m} \tag{8}$$

where r_{dm} is the vector of channels' ratings in DMA d and month m. Equation 8 contains the essence of estimating the distribution of preferences from the ratings data. The left-hand side is an aggregate of the utilities attributable to channels in specific markets and times as a function of observed aggregate ratings. These aggregates vary over time and across markets, as do observed demographics and average bundles. If bundles were assigned randomly to households, the relationship could be consistently estimated as a system of equations by OLS.

Because households select the bundle of channels they purchase, I can not directly estimate the population distribution of β_{ij} from 8. In order to estimate the population distribution, I add co-variates that predict, according to the demand model, the selection. For a given channel, I add a flexible function of the several large channels' penetration rates in that market and the average price of bundles that have the channel in the DMA. The reasoning behind adding selection co-variates is similar to Olley and Pakes (1996). With unlimited computing power, I could use the demand model to match the conditional distribution of ratings. Since this is computationally costly, I instead try to uncover the unconditional distribution from the conditional using a function of the variables that affect the selection according to the demand model. Unlike Olley and Pakes (1996), I do not have a strict result that the covariates I add will always work. Monte Carlo simulations of this procedure at the estimated values suggest that it works with a maximum error of under 5% of the estimated parameters and a median error near zero.

I run the following system of equations regression:

$$r_{dm}\log(r_{dm}T) \quad \approx \quad \widetilde{\beta} + \Pi D_d + \Theta\widetilde{x_m} + \mu_{dm} + \epsilon_{dm} \tag{9}$$

where $\widetilde{\beta}$ is the vector of constant terms[27], μ_{dm} is a vector of variables influencing either the approximation error or the selection into bundles. I use the the channel's DMA penetration rate, the share-weighted average ratings of bundles in the DMA, and the DMA average price of bundles including the channel in the DMA. The residual ϵ_{dm} has a structural interpretation as the DMA's Nielsen sample average of unattributable tastes.

I will now estimate G, the distribution of v_i, using the relation $\epsilon_{dm} = \Upsilon^{dm} v_i$ and semi-parametric assumptions on G. I assume the marginal distributions of G are a mixture of mass points at 0 and exponential distributions with parameters to be estimated. I set the mixture weights equal

[27] $\widetilde{\beta}$ depends on β, but they are not necessarily equal.

to three times the cumulative rating of each channel. I then choose the single parameter of the exponential portion of each distribution so that the implied variance of sample averages from the mixture distribution has the variance of the residuals ϵ_{dm}. I then shift the distributions so that all willingness to pays are positive and the fraction of households which watches the channel for less than ten minutes per week matches the cumulative rating of the channel. Finally, I use a t-copula to fix the covariance structure of G. The copula is based on rank-correlations. It is invariant to CDF and inverse CDF, so it preserves the estimated marginal distributions. The following list summarizes the procedure for estimating v^*_{ij}:

1. Regress ratings on demographics conditional on approximation and selection terms.
2. Parameterize marginal distributions of G as mixtures of mass point at zero and exponential distributions.
3. Choose exponential distribution parameters so that variances of sample averages match variances of residuals from first step.
4. Shift marginal distributions for positive willingness to pay and to match cumulative distributions.
5. Use a copula so that Kendall's τ of G is equal to Kendall's τ of ϵ_{dm} and marginal distributions are preserved.

Given estimates for v^*_{ij}, estimating the remaining demand parameters follows the standard logic of estimating BLP-type demand models. I start with a parameterized distribution of α_i:

$$\alpha_i = \alpha + \pi_p y_i$$

where y_i is household i's income. I form the following moment conditions that identify λ, π_p, and α:

$$\begin{aligned} E[\xi_{jmt}\tilde{z}_{jmt}] &= 0 \\ \xi_{jmt} &= \delta_{jmt}(s_{mt}, x_{mt}, p_{mt}; \hat{\beta}, \hat{\Pi}, \hat{G}, \hat{\Theta}, \sigma_\epsilon, \pi_{ip}, \cdot) - z'_{jmt}\lambda + \alpha p_{jmt} \end{aligned}$$

where $\tilde{Z}$ is a set of instruments that does not include the price of the bundle. I compute $\delta_{jmt}(\cdot)$ using the BLP contraction mapping. As instruments for price, I use the average price of non-competing bundles in near-by markets and the channels available on the bundle. These will be valid instruments if they are uncorrelated with the unobservable portion of mean utility, ξ_{jndm}, but correlated with price. These instruments should be correlated with price through their relationship with costs. I argue further for the validity of these instruments in the empirical identification section. Finally, for efficiency, I also incorporate moment conditions implied by Nash equilibrium in prices which restrict the parameters α and π_p.

B. *Identification and Estimation of Supply: Distributor Costs*

Aggregate input costs, the necessary conditions implied by Nash equilibrium in prices and bundles, and the observed prices and bundles identify input costs. Aggregate input costs are direct evidence. The Nash conditions are indirect evidence; what could input costs have been given the Nash assumption and observed prices and bundles? This section uses the Nash conditions to estimate input costs accounting for factors which are unobservable to the econometrician but known to the distributors at the time of their pricing and bundling decisions.

I parameterize τ_{fc} as a function of channel characteristics $g(c)$ scaled by a function of firm and channel characteristics:

$$\begin{aligned}\tau_{fc} &= g(c)h(f,c)\\ &= (\eta x_c)exp(\varphi z_{fc})\forall c\end{aligned}$$

x_c is a function of a constant term and the Kagan average input cost for channel c. z_{fc} contains firm f's total number of subscribers and whether channel c and whether firm f are vertically integrated. While different channels may have different base rates, I assume the functional form of the effect of distributor size and vertical integration on input cost is the same for all channels. If Comcast has a 30% discount on the base rate of ESPN, it also has a 30% discount on the base rate of CNN, and for any other channel that it is not vertically integrated with.

A weighted average of τ_{fc} over firms predicts an aggregate input cost for each channel c. The Kagan EBCN data set's channel input costs are the empirical counterpart of these averages. One set of moment conditions is simply the model's predicted aggregate input costs should equal observed aggregate input costs: $\{\tau_c\}$.

$$E[\tau_{fc}(\eta,\varphi)] - \tau_c = 0$$

The first order condition to maximize firm f's profits with respect to the price of bundle k in market m is:

$$
\begin{aligned}
\frac{d\Pi_{fm}(\mathbf{b_m}, \mathbf{p_m})}{dp_k} &= \sum_{j \in B_{fm}} (p_j - \sum_{c \in C_j} \tau_{fc}) \frac{dD_{jm}(\mathbf{b_m}, \mathbf{p_m})}{dp_k} + D_{jm}(\mathbf{b_m}, \mathbf{p_m}) \\
&= 0
\end{aligned}
$$

This says that bundle k's optimal price is equal to the input cost of bundle k plus a mark-up that depends on demand conditions and the other bundles in the market. This condition holds in Nash equilibrium for each bundle of each firm, given all other bundles and prices. As demand parameters enter into the mark-up, I use this condition separately to increase efficiency in demand estimation. It plays a more central role in cost estimation, because in its absence the cost parameters are partially identified. In both cases, I directly invert the first order condition to back out implied marginal costs per bundle. I then form moments based on a bundle marginal cost function of bundle characteristics, accounting for the endogenous choice of mark-up by instrumenting with an exogenous predicted mark-up. These moments depend on the price sensitivity of consumers, and thus place extra restrictions on this parameter and improve efficiency of demand. Additionally, a by-product of demand estimation are consistent estimates of the marginal cost per bundle.

To estimate input costs, one could simply project estimated marginal cost per bundle onto the channels included. I do this, but add the aggregate cost moments, the bundling moments, and use the cost parametrization. Explicitly, here are the moments conditions implied by assuming distributors are at a Nash equilibrium in prices:

$$
\begin{aligned}
\epsilon_j &= \hat{mc}_j(\mathbf{p_m}, \mathbf{b_m}, \hat{\beta}, \hat{\Pi}, \hat{G}, \hat{\Theta}, \sigma_\epsilon, \pi_{ip}) - \hat{mc}_j(\eta, \varphi) \\
E[Z'\epsilon] &= 0
\end{aligned}
$$

where the first $\hat{mc}$ are the implied marginal costs per bundle from inverting the price first order necessary condition, and the second $\hat{mc}$ are the aggregate predicted costs per channel. Z is a set of instruments, which contains, in particular, firm size and the extent of vertical integration in that bundle.

I now derive the restrictions from optimal bundling used in estimation. The logic is the same as the use of the optimal pricing conditions. There are only certain cost parameters which satisfy that adding or dropping channels is less profitable than keeping the observed bundles. However, since adding or dropping channels is a discrete choice, the implied restrictions are inequalities. I follow the set-up in Pakes, Porter, Ho, and Ishii (2006).

From the Nash assumption,

$$\Pi_{fm}((\mathbf{b_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p_{fm}}, \mathbf{p_{-fm}})) \quad \geq \quad \Pi_{fm}((\mathbf{b'_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p'_{fm}}, \mathbf{p_{-fm}}))$$

I approximate Π_{fm} using the profits predicted from the model, r_{fm}, which of course depend on input costs.

$$\Pi_{fm}((\mathbf{b_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p_{fm}}, \mathbf{p_{-fm}})) \quad \approx \quad r_{fm}((\mathbf{b_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p_{fm}}, \mathbf{p_{-fm}})) + \nu_{fmb,1} + \nu_{fmb,2}$$

$\nu_{fmb,1}$ is the error in the approximation that is unknown to the firms when making their bundling decision. $\nu_{fmb,1}$ contains measurement error and firm uncertainty. $\nu_{fmb,2}$ is the error in the approximation known to firms at that time. $\nu_{fmb,2}$ contains, for example, the loss a vertically integrated channel would suffer if its integrated distributor carried a competing channel.

Following Pakes, Porter, Ho, and Ishii (2006), I define

$$\Delta\Pi_{fm}(b, b') \quad \equiv \quad \Pi_{fm}((\mathbf{b_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p_{fm}}, \mathbf{p_{-fm}})) - \Pi_{fm}((\mathbf{b'_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p'_{fm}}, \mathbf{p_{-fm}}))$$

and

$$\begin{aligned} \Delta r_{fm}(b, b') \quad &\equiv \quad r_{fm}((\mathbf{b_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p_{fm}}, \mathbf{p_{-fm}})) - r_{fm}((\mathbf{b'_{fm}}, \mathbf{b_{-fm}}), (\mathbf{p'_{fm}}, \mathbf{p_{-fm}})) \\ \nu_{fm,b,b',1} \quad &\equiv \quad \nu_{fmb,2} - \nu_{fmb',2} \\ \nu_{fm,b,b',2} \quad &\equiv \quad \nu_{fmb,2} - \nu_{fmb',2} \end{aligned}$$

I make the following assumption about $\nu_{fm,b,b',2}$.

ASSUMPTION 1: *For two markets m and m' and the same firm,* $\nu_{fm,b,b',2} = \nu_{fm',b,b',2} = \nu_{f,b,b',2}$.

Therefore, any unobservable error in the approximation of profits for adding or dropping channels is common to all markets for a given firm. For example, the benefit of adding Turner Classic Movies, a channel vertically integrated with Time Warner Cable, that is not accounted for in the function Δr is the same in any Time Warner Cable market.

This assumption and the Nash condition imply the optimal bundling moment conditions:

$$E[\Delta r_{fm}(b, b') + \Delta r_{fm'}(b', b)] \geq 0$$

The estimation routine punishes input cost parameters whose implied r functions violate this condition.

C. *Identification and Estimation of Supply: Channel-Distributor Bargaining*

The unobserved parameters of the bargaining game are each channel and distributor's pair-wise bargaining powers ζ_{fc}. I use no additional data in identifying the bargaining powers. They are identified by the estimated cost and demand parameters and the protocol of the bargaining game.

In practice, I choose the values of ζ_{fc} that minimize the distance of the bargaining model's equilibrium input costs and estimated input costs. The demand and pricing model implies a set of input costs which deliver higher profits for both channel and distributor than no agreement. If this set is non-empty, it will usually be an uncountable set. In this case, the two firms will disagree over what point in the set should be chosen. The channel will most often prefer higher input costs, the distributor will always prefer lower input costs. The bargaining model, for a fixed vector of $\zeta_{\mathbf{c}}$, resolves this disagreement. Part of the resolution is due to the bargaining protocol. The rest is due to the bargaining parameters $\zeta_{\mathbf{c}}$. The estimated input costs are an estimate of the actual resolution point. Therefore, the estimated bargaining powers are the $\zeta_{\mathbf{c}}$ which imply equilibrium input costs from the bargaining model as close as possible to estimated input costs.

I use a simplified industry structure in estimation of the bargaining parameters. I assume that the country is served by one large cable provider, one small cable provider, and one satellite provider. The large cable and small cable operate in different markets which only differ in number of households. The satellite provider competes with the cable operators in each market, but it must set the same price and package in both markets. The simplified industry structure reduces the number of players in the bargaining game, which in turn reduces the computational burden of estimation. I take the large cable firm to be Comcast, the small to be an unnamed independent firm, and the satellite firm to be DirecTV. Without a simplification, it would be necessary to solve the bargaining game with many simultaneous negotiations, and to have the downstream competition take place in the thousands of markets across the country. The simplification allows a connection to the estimated cost parameters by having different sized distributors, but without having so many distributors and markets that computing even one equilibrium would take hours.

For estimation, I assume that downstream firms compete in pure bundling. This is weakly less profitable than mixed bundling schemes, but most supply contracts for large channels have Resale

Tier Maintenance which limits the scope for mixed bundling by downstream firms[28].

VI. Estimates and Empirical Identification

I now describe the parameter estimates and their estimated precision. I demonstrate how patterns in the raw data influence the structural estimates. I compare the estimated parameters with previous researchers' estimates of similar parameters.

A. *Demand and Viewership*

Figures 9 and 10 plot simulated marginal distributions of estimated WTP for a subset of channels. Figure 4 plots summary statistics for the simulated marginal distributions of WTP for a subset of channels. I now isolate and analyze the marginal distribution of WTP for BET to demonstrate how empirical patterns in the data manifest in the estimates.

Figure 2 shows ratings for BET in the DMA's where BET has the highest and lowest ratings. BET has its highest ratings in DMA's with relatively high percentages of black households and its lowest ratings in DMA's with relatively low percentages of black households. Statistically, BET's rating co-varies positively with a DMA's percentage of black households. This pattern's structural interpretation is shown in Figure 3; the distribution of willingness to pay for BET by black households is translated to the right, all else equal, according to how the model's transformation of the ratings into willingness to pay co-varies with percentage of black households. However, percentage of black households, or other demographic variables, can not perfectly predict BET's ratings across DMA. I assume the left over variance to come from the variance in tastes that are unattributable to demographics. This final source of variance is what determines the spread in BET's estimated willingness-to-pay after conditioning on demographics.

To estimate the common component of marginal utility of income, α, I match the market shares of bundles of channels predicted by the model with their empirical counterparts. The classic identification problem for this parameter is that unobservable demand shocks affect shares and prices simultaneously. Crawford and Yurukoglu (2008) use an instrumental variables assumption to identify this parameter. I assume that the unobservable portion of utility, net of v^*_{ij} and α_i, δ_j is uncorrelated with the channels in the bundle and average prices of bundles from other markets in the same DMA. This assumption identifies α because prices are correlated with the channels in the bundle and prices of bundles in other markets through costs. By construction, the utility from

[28] Building resale tier maintenance and other discrete contracting clauses would improve the realism of the model. The cost to doing so is sacrificing the convexity of the bargaining problem unless one allows for lotteries over contract clauses.

Figure 2. High and Low Rating DMA's for Black Entertainment Television

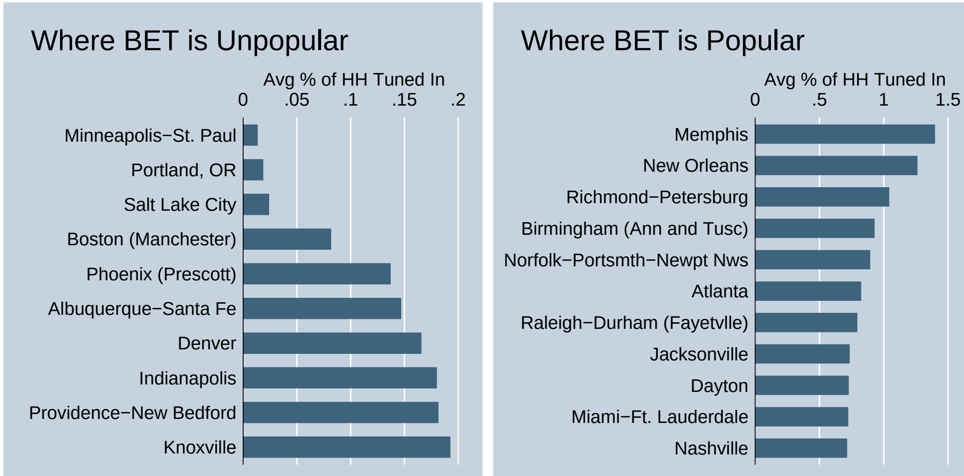
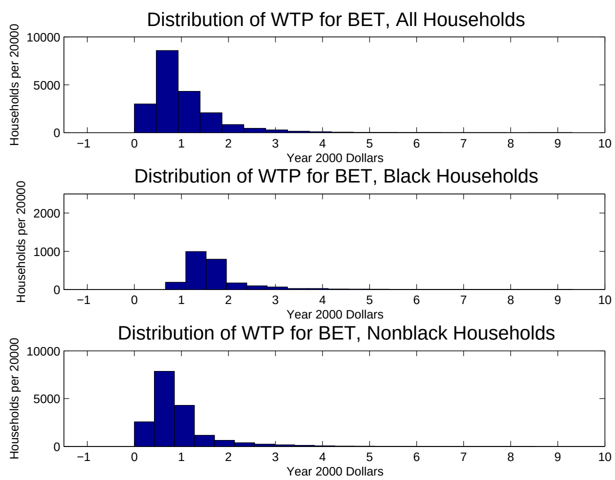


Figure 3. Estimated Distributions of Monthly Willingness to Pay for BET



the channels in the bundle v^*_{ij} is already netted out of δ_j. Prices in other markets are uncorrelated with the unobservable portion of δ_j as long as demand shocks are local. This assumption would not be true if, for example, advertising is DMA or nation-wide. Using both the channels and average prices in other markets over-identifies the model. The results are similar when using only one set of instruments or the other.

Furthermore, I use the moments implied by the pricing equation to increase efficiency. As theory would predict, the instrumental variables assumption generates estimates of the marginal utility of income which imply vastly more reasonable elasticities. Table 5 shows that ignoring the simultaneity of price and market share would produce estimates of the marginal utility of income near

Table 5. Price Sensitivity Parameter Estimates Under Various Moment Restrictions. For the purposes of this table, the estimated standard errors in parentheses do not account for simulation or first stage estimation error.

OLS No Pricing Equation	IV No Pricing Equation	IV Pricing Equation
-0.0396 (.0014)	-0.1412 (.0069)	-0.1552 (.0047)

to zero, with corresponding elasticities near zero as well. The IV assumption and extra moments from the pricing equation improve the credibility of the estimates. The estimated elasticities, under these assumptions, are in Table 6. These estimates of mean own price elasticities, -1.86 for basic service, -5.27 for expanded basic, -12 for digital basic, and -4.3 for satellite, are in line with previous researchers' estimates [29] One improvement of these estimates over previous estimates is that the market specific elasticities may vary depending on what channels are in the bundle and the demographic make-up of the market. Qualitatively, the most notable difference is with the cross-price elasticity estimates of Goolsbee and Petrin (2004). They estimate that the cross price elasticity of the outside good share with respect to the price of premium cable is equal to the cross price elasticity of the expanded basic cable share with respect to the price of premium cable. I estimate, as in Crawford and Yurukoglu (2008), that the former cross price elasticity is near zero. Similarly, Goolsbee and Petrin (2004) estimate that the cross price elasticity of premium cable, basic cable, and the outside good with respect to the price of satellite are almost equal. This paper estimates that these cross price elasticities are much higher for expanded basic and digital cable than for basic cable and the outside good. These estimates reflect the fact that the satellite packages resemble expanded basic and digital basic cable packages.

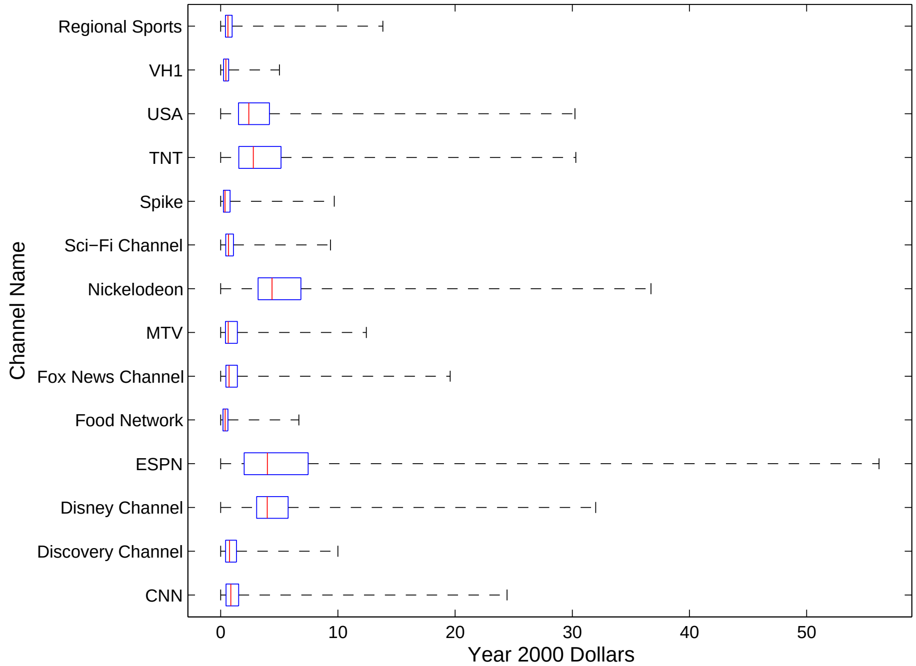
B. *Input Costs*

Table 7 summarizes the estimated marginal costs for bundles of television channels. These estimates are implied by the estimated demand parameters and the distributor's first order condition for choosing prices to maximize profits.

These estimates are combined with the Nash bundling assumption and Kagan average carriage costs per channel to estimate differences in per-channel input costs across distributors. Crawford and Yurukoglu (2008) attempted to project the estimated bundle marginal costs onto the channels in the bundle, but did not find enough variation in the bundles to do so convincingly. By bringing the extra information contained in Kagan's average costs and from the Nash in bundling assumption,

[29] Goolsbee and Petrin (2004) estimate an own-price elasticity of -1.5 for expanded basic, -3.2 for premium cable, and -2.5 for satellite. The FCC (2002) estimated -2.19, the GAO (2004) -3.2, T. R. Beard and Saba (2005) -2.5, and Chipty (2001) -5.9. Chu (2006) and Rennhoff and Serfes (2008) do not report their own-price elasticity estimates.

Figure 4. Box and Whisker Plot of Estimated Monthly Willingness to Pay for a Subset of Channels



I am able to estimate not only channel specific input costs, but also how those input costs differ for downstream firms based on size or vertical integration.

Table 8. Estimated Input Cost Parameters

	Coef	*SE*	*t Statistic*
Constant (η_1)	0.000	0.002	0.000
Kagan Scale (η_2)	1.100	0.023	48.889
Distributor Size (φ_1)	-0.006	0.001	-10.167
Vertical Integration (φ_2)	0.020	0.056	0.358

I present the estimated input cost parameters, η and φ, in Table 8. These estimates imply that Comcast, a distributor with roughly 23 million subscribers, faces input costs 13% below those of a small distributor. The estimated effect of vertical integration is slightly positive, contrary to economic theory, but not statistically significantly different from even large negative values.

The patterns in the data generating these estimates are clear from Tables 9 and 10. Estimated marginal costs and observed prices are lower on average for large distributors, conditional on the characteristics of the bundle. Consequently, I estimate large distributors to have lower per-channel input costs. Similarly, prices and estimated marginal costs for bundles don't vary systematically in a statistically significant way for distributors who offer many of their own vertically integrated channels. One might expect these distributors to at least carry their vertically integrated channels more often than other distributors. This is not true for most of the vertically integrated channels I examine. It is true for some new and small channels that are not part of the analysis. For example, both CNN, a large and highly watched news channel, and CNN International, a smaller channel targeted towards an international audience, were vertically integrated with Time Warner Cable during the sample period. Pricing and carriage decisions for bundles with CNN do not differ systematically for Time Warner Cable compared to other distributors. CNN International, on the other hand, is carried much more often by Time Warner Cable than by other distributors. Table 10 presents statistical evidence to the effect that carriage is not systematically different. More analysis would be necessary to determine whether Time Warner Cable's specific markets have higher tastes for international news, but the pattern holds conditional on market characteristics. Chipty (2001) focuses on a small and specific group vertically integrated channels to find that integration does affect costs and carriage. Here, I show that this is indeed true if one looks at certain less-established channels, but not for the established channels.

Table 9. Regression Analysis of Distributor Size on Price and Estimated Marginal Cost

	Price Regression			Estimated Marginal Cost Regression		
	Coef	*SE*	*t Statistic*	*Coef*	*SE*	*t Statistic*
Distributor Size	-0.059	0.014	-4.070	-0.185	0.030	-6.130
Vertical Integration	-0.073	0.092	-0.790	-0.010	0.194	-0.050
Dummy Variables						
Channels	Yes			Yes		
Year	Yes			Yes		
Tier	Yes			Yes		
Number of Bundles	Yes			Yes		
Year x Tier	Yes			Yes		
Number of Bundles x Tier	Yes			Yes		
N	20117			20117		
R-squared	0.564			0.632		
F(160, 19956)	159.41			211.17		

C. *Bargaining Parameters*

The estimated bargaining parameters are functions of the estimated costs. I find them by searching for the bargaining parameters that produce the estimated input costs as the bargaining game's equilibrium.

The estimates are presented in Table 11. I estimate that bargaining parameters are usually between

Table 10. Carriage of Time Warner Channels by Distributor 2004-2007. CNN and Cartoon Network are each over 15 years old. Boomerang and CNN International are digital channels that began distribution in the 2000's. Carriage for the established channels is not systematically different for the vertically integrated operator Time Warner Cable.

	N	*CNN*	*CNNi*	*Cartoon Network*	*Boomerang*
Charter	1652	0.980	0.078	0.648	0.137
Comcast	2045	0.996	0.007	0.871	0.004
Cox	257	0.988	0.058	0.922	0.144
Time Warner Cable	589	0.988	**0.204**	0.902	**0.447**
Other	6926	0.980	0.008	0.663	0.074

one-half and three-fourths for distributors; cable and satellite distributors are able to capture more of the surplus generated by selling channels to households than if the input costs were the stationary subgame perfect passive beliefs equilibrium of an alternating offers Rubinstein bargaining game. In particular, these estimates reject assuming take-it-or-leave-it offers as the estimated bargaining parameters are neither zero, which would imply channels make take it over leave it offers, nor one, which would imply distributors make take it or leave it offers.

I find that the bargaining parameters are higher for cable firms than satellite firms, even though satellite firms have much larger potential markets than some small cable firms. In equilibrium, satellite firms have lower input costs than small cable firms due to market conditions. This discount would be larger if the two firms had equal bargaining parameters. Within cable firms, large cable firms have higher estimated bargaining parameters than small cable firms for most channels.

VII. Policy Analysis

A. *Socially Optimal Allocation*

Since the social marginal cost of serving an extra household existing programming is zero[30], the static Pareto optimal allocation of multichannel television is for every household to receive the channels it values positively. Market power distorts the equilibrium away from Pareto optimality. Channels are able to negotiate input costs greater than the social marginal cost of zero. Distributors then profit from pricing above their negotiated input cost. Furthermore, the incentives for distributors to separate consumers by bundling could theoretically reduce the efficiency of this market.

[30]Following digital upgrades in the late 1990's, bandwidth congestion had not been an issue for distributors until recently. Distributors now cite high definition (HD) programming and increased data usage as having potential to cause congestion. In that case, positive prices for certain channels could be socially optimal.

I now simulate the effects of regulating bundling on market efficiency and distributional measures by banning distributors from bundling channels together. Stigler (1968), Adams and Yellen (1976), McAfee, McMillan, and Whinston (1989), and other authors have demonstrated the ability of multi-product monopolists to use bundling to increase profits and decrease consumer welfare. By bundling channels together, distributors reduce demand heterogeneity so that they may extract more from households. When distributors must price channels individually, a channel's valuations are dispersed so that to attract new buyers, the channel's price must drop substantially. As long as valuations for channels are not perfectly correlated, the dispersion of valuations for the bundle is lower. A seller can thus set a high price for the bundle and still serve most of the consumers. The theoretical results are parameter dependent; depending on costs and the distribution of preferences for components of the bundle, banning bundling can increase or decrease market efficiency and consumer surplus[31]. I resolve this ambiguity by using this paper's model's estimated parameters to simulate anti-bundling measures.

In addition to assuming the model's parameters invariant to the regulation, I further assume that the set of channels and household preferences do not adjust to banning bundling. These dynamic effects are often mentioned as arguments for why banning bundling would be bad; Channels would exit, or adjust the quality of their programming, and households would not be able to learn about new programming they might enjoy. There are other potential costs and benefits of banning bundling that I do not account for, such as increased accounting and equipment costs for distributors, increased marketing costs for channels, and cognitive effects on consumers.

B. *Baseline Counterfactual Policy Analysis*

I compute the model at the estimated parameter values under two assumptions of downstream competition. Like in estimating bargaining parameters, I compute the model on a simplified national basis with a big cable and small cable competing in separate markets against a nationwide satellite provider[32]. In one case, I assume that the downstream firms engage in pure bundling (PB) where they set prices for just one bundle of all the channels. This case represents the status quo[33]. In the second case, I assume that downstream firms set six prices: three prices corresponding to bundle size in terms of number of channels, and three additional *à la carte* prices. Each firm sets prices for bundle sizes of five, twenty, and sixty channels. Consumers can choose which channels to

[31]This is partly because a firm which bundles may not serve part of the market that values components of the bundle, but not enough to purchase the whole bundle.

[32]As in the case of estimating bargaining parameters, I make this simplification to save computational time. The preferences for individual channels are the estimated preferences. However, the distributors are now fictional, so there are no estimated preferences for their non-channel attributes. I estimate preferences for non-channel attributes so that the simplified model's predictions for aggregate market shares and prices match their empirical counterparts.

[33]Although bundles m

take to fulfill the number for which they paid. This sort of pricing scheme is analyzed and named "Bundle Size Pricing" (BSP) in Chu, Leslie, and Sorensen (2006). In Crawford and Yurukoglu (2008), we studied a downstream model where distributors set prices for individual channels, called component pricing (CP). I also force distributors to set component prices, but only for three large channels: ESPN, The Disney Channel, and Nickelodeon. Therefore, I refer to the simulated equilibria as hybrid BSP-CP equilibra. BSP equilibria are less costly computationally compared to CP equilibria which have more choice variables. This saving is important in models with endogenous input costs because the downstream equilibria are nested in the upstream equilibrium and must be computed many times.

Table 12. Simulated welfare results when distributors must offer BSP for bundles sizes of 5, 20, and 60, and CP for The Disney Channel, ESPN, and Nickelodeon. BSP-CP Naive refers to holding input costs fixed at their pre-regulation values. The BSP-CP column are the results after renegotiation of input costs. The units are year 2000 dollars per month per television household, of which there are roughly 115 million.

	Bundling	BSP-CP Naive	*Change*	BSP-CP	*Change*
Total Welfare	68.5072	68.1347	-0.54%	67.127	-2.015%
Mean CS	33.5253	35.553	6.05%	33.2036	-0.960%
25th Prctile CS	20.2475	22.3289	10.28%	20.0889	-0.783%
Median CS	31.0705	33.2679	7.07%	30.8137	-0.827%
75th Prctile CS	43.866	46.1534	5.21%	43.6112	-0.581%
Wired Households	0.8991	0.9155	1.82%	0.8997	0.067%
Mean Expenditure	34.9819	32.5816	-6.86%	33.9235	-3.026%
Mean # of Channels	77.3218	53.5139	-30.79%	50.3085	-34.936%
Big Cable Profits	12.8978	13.2481	2.72%	13.0029	0.815%
Small Cable Profits	2.3518	2.4677	4.93%	2.3906	1.650%
Satellite Profits	3.9786	4.2643	7.18%	4.1198	3.549%
Channel Profits	15.7538	12.6015	-20.01%	14.4101	-8.529%
Total Industry Profits	34.982	32.582	-6.86%	33.923	-3.026%

The results are in Table 12. The BSP-CP hybrid equilibria has about 2% lower total welfare, 1% lower average consumer surplus, and 3% lower industry profits than the PB equilibrium. If input costs did not re-equilibrate conditional on downstream BSP-CP (like the pure CP equilibrium examined in Crawford and Yurukoglu (2008)), one would wrongly estimate consumer surplus to increase by 6%. Within consumers, there are some whose welfare increases in the BSP-CP equilibrium. These are households which value the à la carte channels less than their new equilibrium price. These households sacrifice the utility from those channels, but pay enough of a lower bill that they benefit. Low income households fare slightly better than high income households as those who pay a lower bill value the extra income more highly. This pattern would be stronger if the channels offered individually had higher estimated willingness to pay for the channels for high income households.

Input costs re-equilibrate to a higher level when downstream firms engage in BSP-CP. This is because forcing BSP-CP instead of PB alters the marginal surpluses generated by channels and distributors. Moving from PB to BSP-CP changes the nature of the distributor's pricing problem. Under PB, demand is driven by valuations for the whole bundle. Whether this valuation is high because a household values one channel very highly or many channels moderately is of no importance. The Nash bargaining solution prefers lower input costs with higher quantity. With BSP-CP, the Nash bargaining solution prefers higher prices and lower quantity.

Figures 5, 6, and 7 provide graphical examples to engender intuition. Figure 5 depicts Adams and Yellen (1976). Consider two goods with dispersed valuations and fixed marginal costs of zero. Pricing each good individually would require the seller to choose an intermediate price. The seller would miss out on the surplus enjoyed by high valuation consumers and on the low valuations above marginal cost which it does not serve. However, as long as valuations between the two goods are not perfectly correlated, the valuation of the bundle will be less dispersed. In 5, I chose underlying valuations that are highly negatively correlated to emphasize this point. Pricing only the bundle allows the seller to capture most of the combined surplus. Forgetting bundling for a moment, consider the determination of input costs for a single good in a bilateral monopoly with linear fee contracts in Figure 6. For a given input cost from the y-axis of the left panel, the seller in the right panel maximizes profit by choosing price to equate marginal revenue and marginal cost. The area of the upper producer surplus rectangle is the downstream seller's profit. The area of the lower producer surplus rectangle is the upstream producer's profit. The Nash product in the left panel is a weighted[34] geometric average of these two profits. The equilibrium input cost maximizes the Nash product. Figure 7 combines these two examples for the determination of input costs in a situation analogous to bundling versus component pricing. It repeats 6 for two goods which have the same underlying mean valuations, but different dispersions. Associating the demand curve for the less dispersed valuations with demand for the bundle, one can see that the equilibrium input cost for this good is lower than for the good with highly dispersed valuations.

Finally, I repeat the left panel of Figure 7 using estimated valuations. Figure 8 depicts the graph of the Nash products for ESPN and a monopoly distributor that has two goods, ESPN and the group of all channels besides ESPN. In the case of bundling, the Nash product is higher except at high input costs. However, the Nash product is maximized at a higher input cost when the distributor offers the two goods individually. Households with a low valuation for ESPN opt-out of purchasing ESPN. Channels and distributors jointly try to extract from high valuation consumers instead.

[34]In this case, I use equal weights. In the general model, ζ_{fc} is the weighting for each pair fc.

Figure 5. The dispersion in valuations for the bundle will be lower than the dispersion in valuations of the components.

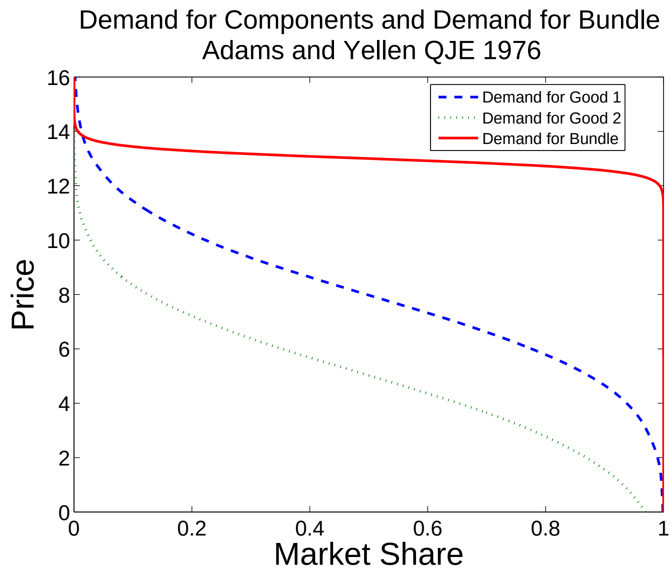


Figure 6.

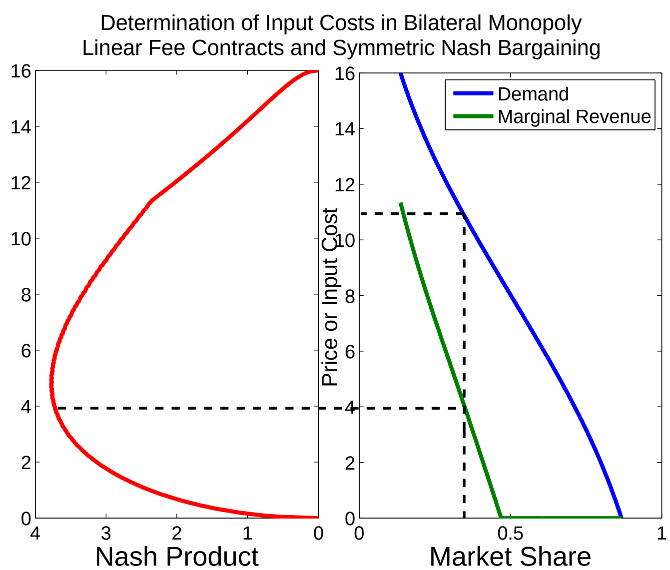


Figure 7. Change in graph of Nash products for goods with more or less dispersed valuations.

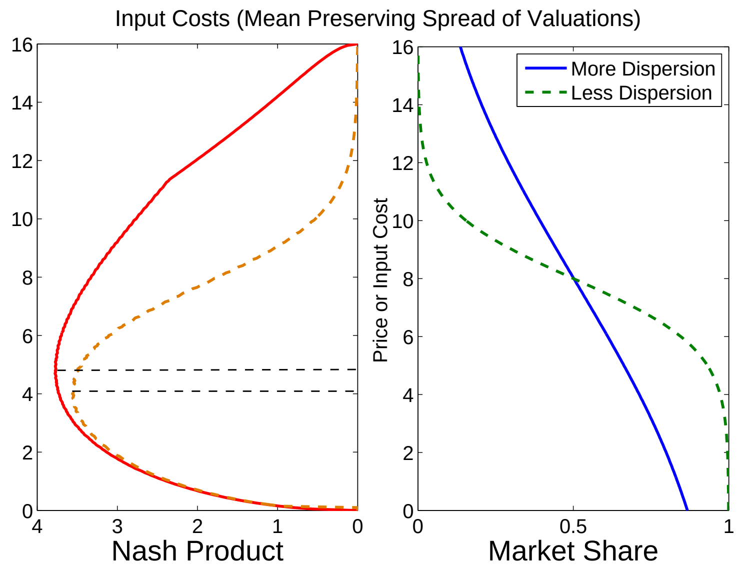
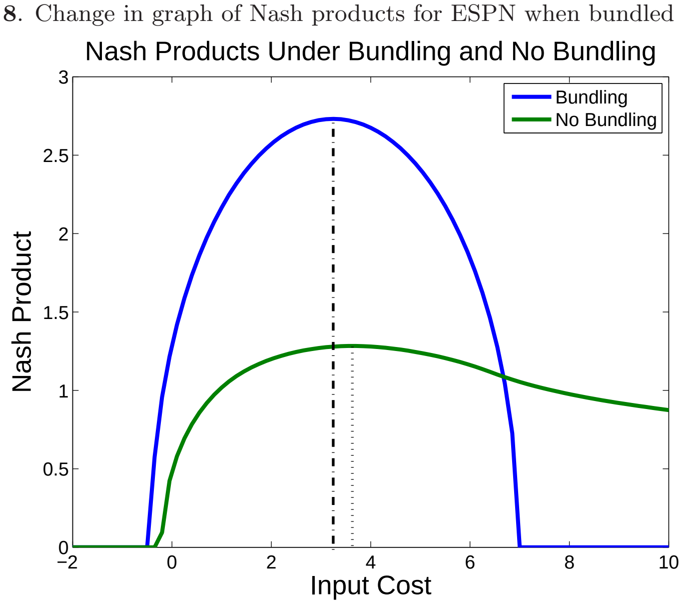


Figure 8. Change in graph of Nash products for ESPN when bundled and not.



VIII. Conclusion

This paper models the US multichannel television industry in order to simulate the welfare effects of banning the bundling of television channels by distributors of multichannel television service. I allow for the re-equilibration of input costs following anti-bundling regulations. I find that banning bundling and forcing a hybrid of bundle size and component pricing is likely to decrease welfare for most of society. Assuming that input costs do not change would predict beneficial effects for consumers relative to the model of this paper. Under this paper's bargaining model and parameter estimates, equilibrium input costs rise when bundling is banned. High equilibrium input costs are partially passed on to consumers. Most consumers are ultimately hurt by anti-bundling regulations.

Further research on the dynamic incentives facing consumers and firms would add credibility to the predictions presented here. In particular, how might channels respond other than re-negotiating input costs? Would some channels exit, would new channels enter, and what would happen to the quality of programming offered and advertising levels? How would consumers formation of preferences over programming change when they receive fewer channels? Extending the demand, pricing, and bargaining model of this paper to account for dynamic reactions of consumers and firms would increase the credibility of predictions about regulation in this industry.

Modeling supply-side dynamics would also enhance the theory behind the bargaining model. In this paper, bargaining parameters are structural parameters; there is no explicit theory of where they come from. They are assumed invariant to the policy change. I conjecture they are partially determined by the ability of either side of the market to replace the other, by backward or forward integration as in Katz (1987). More theory and empirical evidence about what determines the bargaining parameters is necessary.

References

Adams, W., and J. Yellen (1976): "Commodity Bundling and the Burden of Monopoly," *Quarterly Journal of Economics*, 90(3), 475–98.

Berry, S. (1994): "Estimating Discrete Choice Models of Product Differentiation," *Rand Journal of Economics*, 25(2), 242–262.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Binmore, K. (1987): "Perfect Equilibria in Bargaining Models," in *The Economics of Bargaining*, ed. by K. Binmore, and P. Dasgupta.

Chipty, T. (2001): "Vertical Integration, Market Foreclosure, and Consumer Welfare in the Cable Television Industry," *American Economic Review*, 91(3), 428–453.

Chipty, T., and C. Snyder (1999): "The Role of Firm Size in Bilateral Bargaining: A Study of the Cable Television Industry," *Review of Economics and Statistics*, 31(2), 326–340.

Chu, C. S. (2006): "The Effects of Satellite Entry on Product Quality for Cable Television," mimeo, Stanford University.

Chu, C. S., P. Leslie, and A. Sorensen (2006): "Incomplex Alternatives to Mixed Bundling," mimeo, Stanford University.

Crawford, G. (2000): "The Impact of the 1992 Cable Act on Household Demand and Welfare," *RAND Journal of Economics*, 31(3), 422–449.

——— (2006): "Cable Television Regulation in the Satellite Era," Working Paper, University of Arizona.

Crawford, G., and M. Shum (2006): "Empirical Modeling of Endogenous Quality Choice: The Case of Cable Television," mimeo, University of Arizona.

Crawford, G., and A. Yurukoglu (2008): "The Welfare Effects of Bundling in Multichannel Television," mimeo, University of Arizona.

de Fontenay, C., and J. Gans (2007): "Bilateral Bargaining with Externalities," Available http://works.bepress.com/joshuagans/14.

FCC (2002): "2001 Report on Cable Industry Prices," Discussion paper, Federal Communications Commission, Available at http://www.fcc.gov/mb/csrptpg.html.

——— (2004): "Report on the Packaging and Sale of Video Programming to the Public," Discussion paper, Federal Communications Commission, November 18, 2004. Available at http://www.fcc.gov/mb/csrptpg.html.

——— (2005): "Eleventh Annual Report on the Status of Competition in the Market for the Delivery of Video Programming (2004 Report)," Discussion paper, Federal Communications Commission, Available at http://www.fcc.gov/mb/csrptpg.html.

——— (2006): "Further Report on the Packaging and Sale of Video Programming to the Public," Discussion paper, Federal Communications Commission, February, 2006. Available at http://www.fcc.gov/mb/csrptpg.html.

GAO (2004): "Telecommunications: Wire-Based Competition Benefited Consumers in Selected Markets," Discussion paper, General Accounting Office.

Goolsbee, A., and A. Petrin (2004): "Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," *Econometrica*, 72(2), 351–81.

Hart, O., and J. Tirole (1990): "Vertical Integration and Market Foreclosure," *Brookings papers on Economic Activity: Microeconomics*, pp. 205–285.

Hausman, J., and G. Leonard (1997): "Superstars in the National Basketball Association: Economic Value and Policy," *Journal of Labor Economics*, 15, 586–624.

Ho, K. (2008): "Insurer-Provider Networks in the Medical Care Market," American Economic Review, forthcoming.

Horn, H., and A. Wolinsky (1988): "Bilateral Monopoly and Incentives for Merger," *The RAND Journal of Economics*, 19, 408–419.

Katz, M. L. (1987): "The Welfare Effects of Third-Degree Price Discrimination in Intermediate Good Markets," *The American Economic Review*, 77(1), 154–167.

McAfee, R. P., J. McMillan, and M. Whinston (1989): "Multiproduct Monopoly, Commodity Bundling, and Correlation of Values," *Quarterly Journal of Economics*, 104(2), 371–383.

McAfee, R. P., and M. Schwartz (1994): "Opportunism in Multilateral Vertical Contracting: Nondiscrimination, Exclusivity, and Uniformity," *The American Economic Review*, 84(1), 210–230.

Muthoo, A. (1999): *Bargaining Theory with Applications*. Cambridge University Press, Cambridge.

Nash, J. F. (1950): "The Bargaining Problem," *Econometrica*, 18(2), 155–162.

Olley, S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64, 1263–1297.

Pakes, A., J. Porter, K. Ho, and J. Ishii (2006): "Moment Inequalities and Their Application," mimeo, Harvard University.

Prat, A., and A. Rustichini (2003): "Games Played through Agents," *Econometrica*, 71(4), 989–1026.

Raskovich, A. (2003): "Pivotal Buyers and Bargaining Position," *Journal of Industrial Economics*, 51(4), 405–426.

Rennhoff, A. D., and K. Serfes (2008): "Estimating the Effects of a la Carte Pricing: The Case of Cable Television," *SSRN eLibrary*.

Rey, P., and T. Verge (2004): "Bilateral Control with Vertical Contracts," *The RAND Journal of Economics*, 35(4), 728–746.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50(1), 97–109.

Segal, I., and M. D. Whinston (2003): "Robust Predictions for Bilateral Contracting with Externalities," *Econometrica*, 71(3), 757–791.

Seim, K. (2006): "An Empirical Model of Firm Entry with Endogenous Product-Type Choices," *RAND Journal of Economics*, 37(3), 619–640.

Stigler, G. (1968): "A Note on Block Booking," in *The Organization of Industry*, ed. by G. Stigler. Richard D. Irwin.

T. R. Beard, G. S. Ford, R. C. H., and R. P. Saba (2005): "Fragmented Duopoly: A Conceptual and Empirical Investigation," *Journal of Business*, 78.

A. Channel Distributor Negotiations with Renegotiation

This appendix recasts the bargaining game with an alternative assumption about what happens when a pair disagrees. After all pairs conclude their meetings, the set of agreements are observed by all participants. Pairs which do not reach agreement are forbidden from meeting again, and the agreeing pairs re-negotiate conditional on all the disagreements. Play continues until no new disagreeing pairs emerge after a round of negotiations.

The timing of the game is:

All pairs start active.

Stage 1 Active channel and distributor pairs meet bilaterally to decide input costs.

Stage 2 Disagreements are announced. Disagreeing pairs are no longer active.

Repeat Stage 1 and 2 until no more disagreements.

Stage 3 Distributors compete downstream.

Let $\Psi = \{\tau_{fc}\}$ be a set of input costs, a scalar for each pair of distributors and channels. If there is no agreement between a distributor and a channel, then the input cost is positive infinity. Ψ is decided in Stage 1. In each bilateral meeting, τ_{fc} solves:

$$\max_{\tau_{fc}} \left[\Pi_f(\tau_{fc}; \widehat{\Psi^S_{-fc}}) - \Pi_f(\infty; \widehat{\Psi^{S\setminus\{fc\}}_{-fc}})\right]^{\zeta_{fc}} \left[\Pi_c(\tau_{fc}; \widehat{\Psi^S_{-fc}}) - \Pi_c(\infty; \widehat{\Psi^{S\setminus\{fc\}}_{-fc}})\right]^{1-\zeta_{fc}}$$

given $\widehat{\Psi^S_{-fc}}$ which is a candidate set of all other input costs when S is the set of active pairs. In a meeting, the channel and the distributor share $\widehat{\Psi^S_{-fc}}$ for all S.

The Perfect Nash equilibrium to this game consists of a set of agreeing pairs, $S^\star$, a set of input costs for each possible subset of pairs, Ψ^S, $S \in 2^{S^\star}$, and conditional downstream prices, such that each $\tau_{fc} \in \Psi^{S^\star}$ is the outcome of Nash bargaining given the other input costs $\Psi^{S^\star}_{-fc}$ and disagreement values implied by Ψ^S, $S \in 2^{S^\star}$ and downstream competition.

I solve the equilibrium numerically. Starting with the case where all negotiations end in disagreement, I calculate payoffs for each player. These serve as the disagreement values for the set of negotiations where only one pair may agree. This set in turn serves as disagreement points for the set of negotiations where exactly two pairs may agree. Continuing up until all possible disagreement points are found, I then fix an initial set of input costs and an arbitrary ordering of the pairs.

I compute the Nash bargaining solution for each pair given the set of input costs and the known disagreement points, updating the set of input costs after each negotiation. A Nash equilibrium is found when the set of input costs does not change from round to round.

A pure strategy Nash equilibrium does not always exist due to the discontinuity of payoffs under breakdown. It is also possible that there are multiple equilibria. I test for multiple equilibria by choosing different starting points. I do not ever find multiple equilibria.

Table 1. Sample Statistics, Bundle Purchase Data

Variable	Nobs	Mean	SDev	Min	Max
Market Types					
Basic Only	20,117	0.601	0.49	0.00	1.00
Basic + Exp. Basic	20,117	0.319	0.47	0.00	1.00
Basic + Dig. Basic	20,117	0.034	0.18	0.00	1.00
Basic + Exp. Basic + Dig. Basic	20,117	0.045	0.21	0.00	1.00
All Markets					
Price	20,117	$23.75	$9.27	$1.82	$117.13
Local Market Share	20,117	0.461	0.259	0.010	0.990
Total Cable Channels	20,117	20.0	15.6	0.0	176.0
Basic Only Markets					
Basic Service					
Price	12,105	$24.14	$6.07	$1.94	$80.47
Local Market Share	12,105	0.551	0.209	0.010	0.990
Total Cable Channels	12,105	17.4	9.3	0.0	95.0
Basic + Exp. Basic Markets					
Basic Service					
Price	3,188	$13.22	$5.34	$1.82	$47.84
Local Market Share	3,188	0.123	0.158	0.010	0.889
Total Cable Channels	3,188	8.1	6.9	0.0	49.0
Exp. Basic Service					
Price	3,188	$27.23	$7.23	$4.98	$71.73
Local Market Share	3,188	0.559	0.193	0.010	0.969
Total Cable Channels	3,188	26.9	9.8	3.0	84.0
Basic + Dig. Basic Markets					
Basic Service					
Price	334	$29.32	$8.58	$3.19	$50.34
Local Market Share	334	0.517	0.183	0.029	0.924
Total Cable Channels	334	41.4	13.2	2.0	66.0
Dig. Basic Service					
Price	334	$45.97	$14.63	$8.29	$113.10
Local Market Share	334	0.120	0.081	0.010	0.705
Total Cable Channels	334	70.0	16.5	33.0	124.0
Basic + Exp. Basic + Dig. Basic Markets					
Basic Service					
Price	300	$13.37	$5.54	$5.18	$38.75
Local Market Share	300	0.220	0.119	0.011	0.625
Total Cable Channels	300	7.6	5.5	1.0	35.0
Exp. Basic Service					
Price	300	$36.24	$8.74	$13.35	$71.73
Local Market Share	300	0.367	0.145	0.013	0.799
Total Cable Channels	300	47.0	10.8	19.0	89.0
Dig. Basic Service					
Price	300	$48.33	$13.74	$18.63	$117.13
Local Market Share	300	0.124	0.077	0.010	0.474
Total Cable Channels	300	81.2	20.5	39.0	176.0

Table 2. Real Price and Total Channels of DirecTV Total Choice Bundle by Year

DirecTV Total Choice Bundle			
	Year	*Real Price*	*Channels*
	1997	32.18	37
	1998	31.68	37
	1999	31.00	37
	2000	29.99	37
	2001	31.10	80
	2002	30.62	85
	2003	31.81	87
	2004	36.45	86
	2005	35.26	86
	2006	35.87	109
	2007	40.14	95

Table 4. Sample Statistics, Ratings Data, Selected Networks

Network	Nobs	Mean	SDev	Min	Max
ABC Family	1,482	0.42	0.13	0.05	0.94
AMC	1,482	0.52	0.16	0.12	1.31
BET	1,477	0.43	0.32	0.01	2.38
Bravo	1,472	0.25	0.16	0.01	0.86
CNN	1,481	0.75	0.32	0.21	2.82
Comedy	1,482	0.49	0.18	0.09	1.41
CMT	1,467	0.19	0.13	0.01	0.90
Disney	1,482	1.19	0.42	0.13	2.99
ESPN	1,482	0.91	0.45	0.17	3.68
Food	1,481	0.41	0.20	0.01	1.12
Lifetime	1,563	0.90	0.37	0.01	2.19
MTV	1,482	0.70	0.23	0.10	1.79
Natl. Geog.	1,109	0.10	0.08	0.00	0.53
SoapNet	1,210	0.11	0.11	0.00	0.70
SPEED	1,037	0.12	0.09	0.00	0.62
USA	1,481	1.17	0.36	0.17	2.57
VH1	1,480	0.36	0.13	0.03	0.96
Weather	1,478	0.30	0.21	0.01	2.69

Year	*Variable*	*N*	*Mean*	*Std. Dev.*	*Min*	*Max*
1998	Subscribers (Millions)	84	28.65	27.40	0.3	74.9
	Average License Fee ($)	83	0.12	0.22	0	1.7
	Net Advertising Revenue (Millions of $)	84	72.70	137.64	0	655.1
	Total Expenses (Millions of $)	83	89.94	128.36	0.2	596.6
1999	Subscribers (Millions)	91	29.97	28.57	0.1	77.6
	Average License Fee ($)	91	0.11	0.19	0	1.2
	Net Advertising Revenue (Millions of $)	91	82.44	156.03	0	776.3
	Total Expenses (Millions of $)	90	98.39	138.48	0.3	644.2
2000	Subscribers (Millions)	100	31.18	29.33	0	80
	Average License Fee ($)	100	0.11	0.18	0	1.14
	Net Advertising Revenue (Millions of $)	100	89.09	165.71	0	825
	Total Expenses (Millions of $)	99	103.15	151.76	0.3	786.2
2001	Subscribers (Millions)	103	36.13	30.66	0	84.2
	Average License Fee ($)	103	0.12	0.18	0	1.3
	Net Advertising Revenue (Millions of $)	103	84.98	155.27	0	709.5
	Total Expenses (Millions of $)	102	110.13	160.57	0.4	882.5
2002	Subscribers (Millions)	109	39.84	31.72	0	87.6
	Average License Fee ($)	109	0.12	0.20	0	1.6
	Net Advertising Revenue (Millions of $)	109	83.47	149.95	0	681.1
	Total Expenses (Millions of $)	108	115.09	169.95	0.5	914.8
2003	Subscribers (Millions)	114	42.19	31.87	0	88.1
	Average License Fee ($)	114	0.13	0.23	0	1.93
	Net Advertising Revenue (Millions of $)	114	91.37	160.41	0	725.4
	Total Expenses (Millions of $)	112	113.69	160.47	0.9	932.2
2004	Subscribers (Millions)	123	42.71	32.23	0	89
	Average License Fee ($)	123	0.14	0.25	0	2.28
	Net Advertising Revenue (Millions of $)	123	94.72	169.32	0	798.7
	Total Expenses (Millions of $)	120	108.44	145.13	0.2	789.5
2005	Subscribers (Millions)	136	42.37	32.83	0	90.9
	Average License Fee ($)	136	0.14	0.27	0	2.6
	Net Advertising Revenue (Millions of $)	136	97.33	176.02	0	880.4
	Total Expenses (Millions of $)	133	109.54	148.30	3	831.3
2006	Subscribers (Millions)	137	46.21	32.63	0.2	94.1
	Average License Fee ($)	137	0.15	0.30	0	2.91
	Net Advertising Revenue (Millions of $)	136	103.75	178.03	0	925.1
	Total Expenses (Millions of $)	134	119.71	158.45	4.1	881.8

Table 3. Summary Statistics from EBCN

Figure 9. Marginal Distributions of Estimated Monthly Willingness to Pay for a Subset of Channels

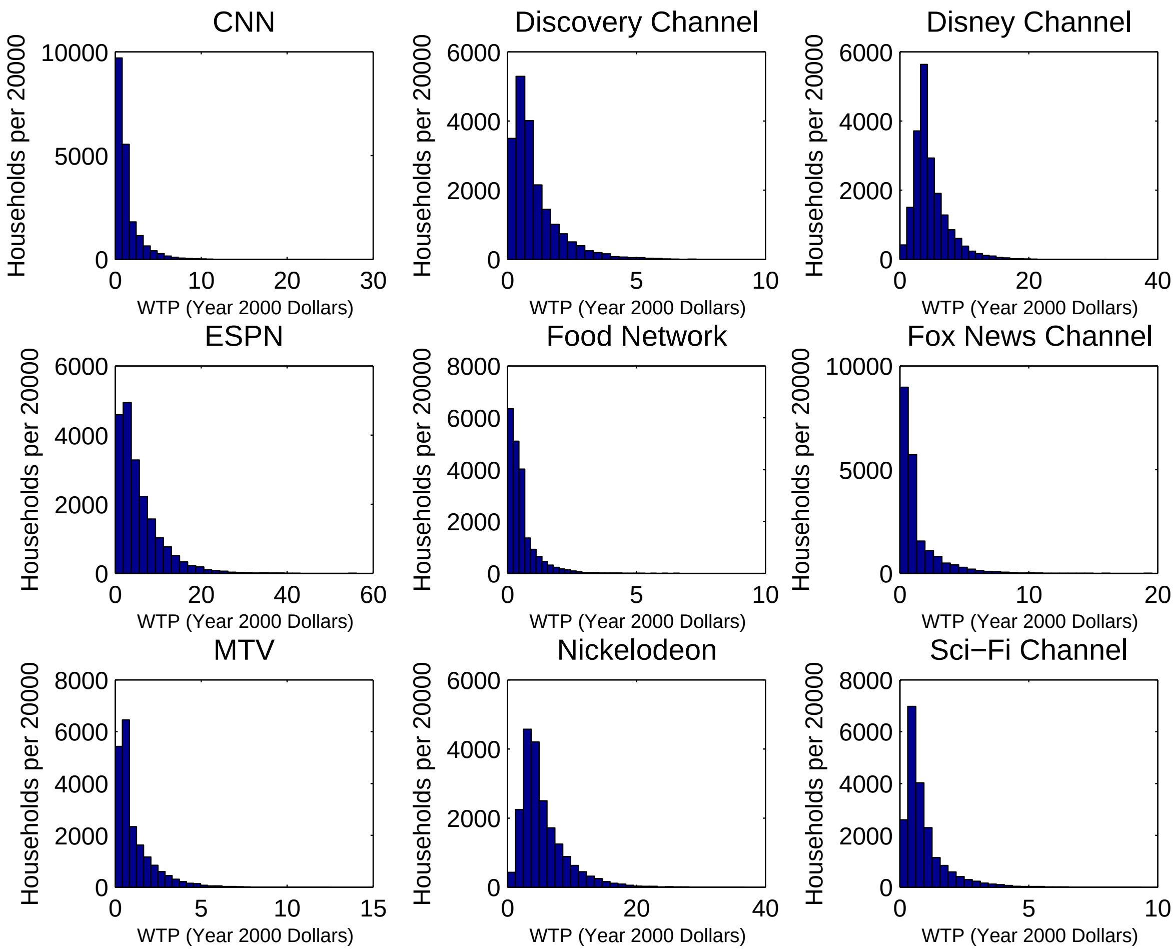


Figure 10. Marginal Distributions of Estimated Monthly Willingness to Pay for a Subset of Channels

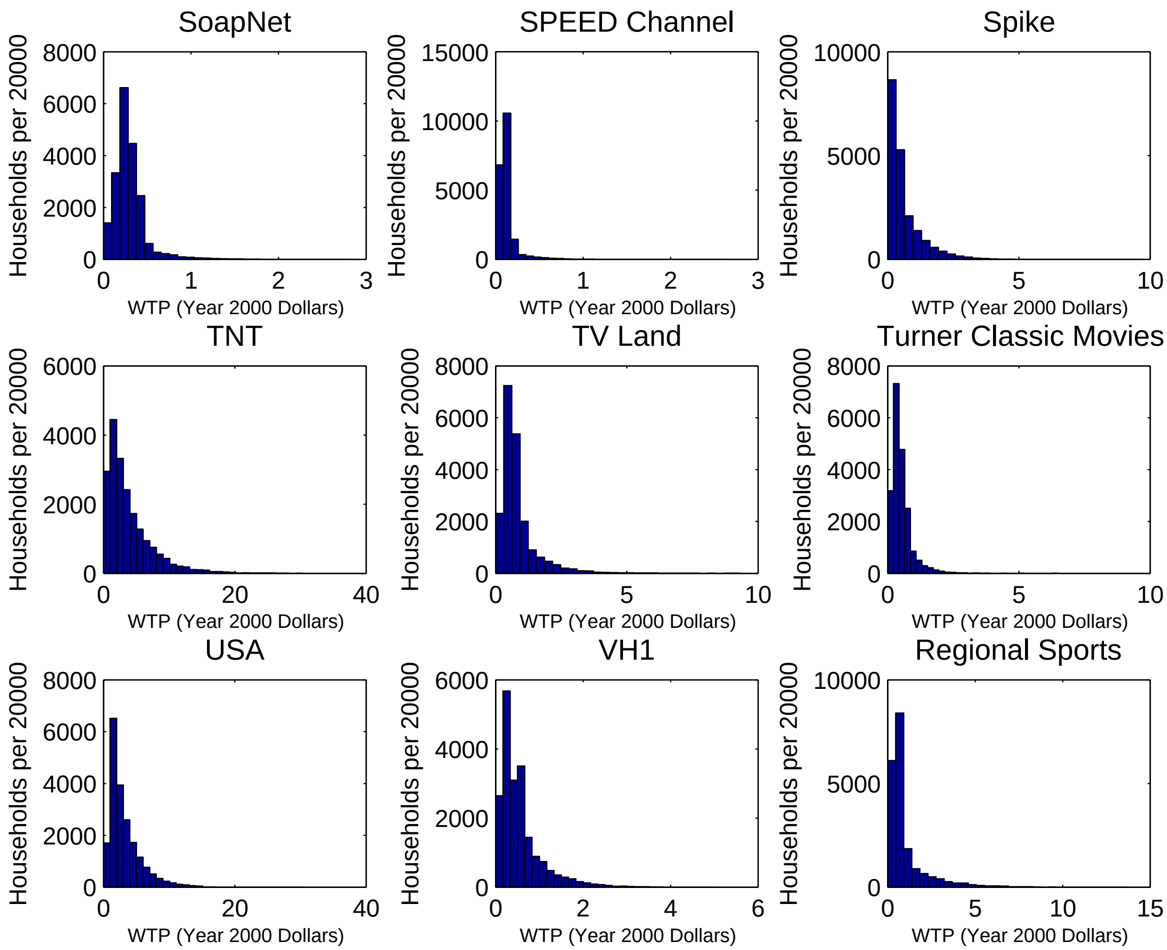


Table 6. Estimated Price Elasticities in Markets where Cable Operator Offers Basic, Expanded Basic, and Digital Basic Bundles

Price Elasticity of	wrt	Mean	Std. Dev.
Basic	Outside Good	0.1968	0.3407
	Basic	-1.861	1.391
	Expanded Basic	1.524	1.214
	Digital Basic	0.543	0.551
	Satellite	0.471	0.719
Expanded Basic	Outside Good	0.156	1.880
	Basic	0.316	0.742
	Expanded Basic	-5.276	2.334
	Digital Basic	2.140	1.976
	Satellite	1.339	1.420
Digital Basic	Outside Good	0.037	0.090
	Basic	0.273	0.722
	Expanded Basic	5.658	2.463
	Digital Basic	-12.077	3.510
	Satellite	1.717	1.486
Satellite	Outside Good	0.036	0.211
	Basic	0.144	0.700
	Expanded Basic	2.004	1.808
	Digital Basic	1.137	1.198
	Satellite	-4.326	2.247

Table 7. Summary of Estimated Marginal Costs, by Bundle Type

	Median Price	Median MC	Std. Dev. MC
All	23.818	8.492	21.096
Basic	22.688	8.172	19.602
Expanded Basic	34.595	8.127	23.139
Digital Basic	46.847	26.811	31.792

Table 11. Distributors' bargaining powers for various channels. Channels' bargaining powers are one minus the estimate.

Estimated Bargaining Powers			
Channel	Big Cable	Small Cable	Satellite
ABC Family Channel	0.60	0.57	0.58
American Movie Classics (AMC)	0.65	0.63	0.63
Animal Planet	0.56	0.56	0.54
Arts & Entertainment (A&E)	0.75	0.74	0.73
Black Entertainment Television (BET)	0.76	0.75	0.75
Bravo	0.39	0.61	0.36
CNBC	0.05	0.08	0.00
CNN	0.62	0.59	0.56
Cartoon Network	0.92	0.93	0.91
Comedy Central	0.71	0.70	0.68
Country Music TV (CMT)	0.56	0.55	0.55
Court TV	0.76	0.76	0.75
Discovery Channel	0.68	0.67	0.66
Disney Channel	0.67	0.64	0.62
E! Entertainment Television	0.35	0.31	0.31
ESPN	0.54	0.54	0.40
ESPN 2	0.47	0.43	0.43
ESPN Classic Sports	0.02	0.00	0.57
FX	0.67	0.65	0.64
FitTV	0.06	0.02	0.00
Food Network	0.72	0.73	0.71
Fox News Channel	0.72	0.71	0.68
Fox Soccer Channel	0.00	0.00	0.00
Fuel	0.60	0.59	0.56
Fuse	0.52	0.61	0.61
Game Show network	0.70	0.70	0.69
GalaVision	0.34	0.32	0.32
Golf Channel	0.00	0.00	0.00
Great American Country	0.10	0.09	0.10
HGTV	0.63	0.63	0.61
Hallmark Channel	0.74	0.74	0.72
History Channel	0.70	0.68	0.67
Independent Film Channel (IFC)	0.00	0.00	0.00
Lifetime	0.87	0.86	0.85
MSNBC	0.55	0.54	0.53
MTV	0.69	0.68	0.65
MTV2	0.12	0.12	0.12
NFL Network	0.00	0.00	0.00
National Geographic Channel	0.02	0.00	0.04
Nickelodeon	0.91	0.91	0.89
Noggin / The N	0.20	0.10	0.13
Oxygen	0.30	0.26	0.27
Sci-Fi Channel	0.72	0.72	0.71
SoapNet	0.48	0.46	0.47
SPEED Channel	0.01	0.04	0.03
Spike	0.60	0.58	0.56
TBS	0.88	0.88	0.85
TNT	0.76	0.76	0.70
TV Guide Channel	0.64	0.64	0.62
TV Land	0.76	0.76	0.75
TLC (The Learning Channel)	0.68	0.67	0.66
Toon Disney	0.79	0.79	0.78
Travel Channel	0.57	0.56	0.56
Turner Classic Movies	0.42	0.38	0.39
USA	0.83	0.83	0.80
Versus	0.05	0.00	0.00
VH1	0.64	0.62	0.62
WE: Womens Entertainment	0.18	0.14	0.15
The Weather Channel	0.67	0.66	0.65
Regional Sports	0.15	0.10	0.02

Table 13. Estimated input costs pre- and post- anti-bundling regulation.

Input Cost Renegotiation							
Channel	Bundling Estimated Input Cost			BSP Estimated Input Cost			Mean Change
	Big Cable	*Small Cable*	*Satellite*	*Big Cable*	*Small Cable*	*Satellite*	
ABC Family	0.215	0.245	0.220	0.229	0.261	0.232	6.2%
Animal Planet	0.078	0.088	0.084	0.118	0.139	0.120	50.5%
BET	0.137	0.154	0.140	0.140	0.161	0.144	3.2%
CNN	0.424	0.489	0.446	0.481	0.564	0.510	14.5%
Discovery	0.263	0.294	0.270	0.304	0.348	0.304	15.5%
Disney Channel	1.647	1.872	1.716	1.962	2.188	2.029	18.1%
ESPN	2.521	2.862	2.623	3.225	3.478	3.409	26.5%
ESPN2	0.224	0.255	0.230	0.285	0.337	0.292	28.8%
FX	0.273	0.311	0.287	0.324	0.376	0.341	19.4%
Fox News	0.244	0.274	0.250	0.282	0.320	0.285	15.4%
History Channel	0.176	0.198	0.180	0.214	0.244	0.222	22.7%
MTV	0.273	0.304	0.284	0.293	0.367	0.301	11.5%
Nickelodeon	0.391	0.443	0.406	0.497	0.553	0.512	26.0%
TBS	0.283	0.321	0.290	0.286	0.333	0.290	1.6%
TNT	0.833	0.947	0.865	0.849	0.980	0.866	1.8%
TCM	0.273	0.311	0.284	0.332	0.387	0.346	22.6%
USA	0.436	0.493	0.456	0.433	0.487	0.450	-1.0%
Regional Sports	1.067	1.209	1.114	1.097	1.361	1.131	5.6%
Total	9.327	10.584	9.699	10.865	12.322	11.287	16.4%