The Welfare Effects of Bundling in Multichannel Television Markets *

Gregory S. Crawford† Ali Yurukoglu‡

April 2011

Abstract

We measure how the bundling of television channels affects short-run social welfare. We estimate an industry model of viewership, demand, pricing, bundling, and input market bargaining using data on ratings, purchases, prices, bundle composition, and aggregate input costs. We conduct counterfactual simulations of à la carte policies that require distributors to offer individual channels for sale to consumers. We estimate that negotiated input costs rise by 103.0% in equilibrium under à la carte. These are passed on as higher prices, offsetting consumer surplus benefits from purchasing individual channels. Mean consumer and total surplus change by an estimated -5.4% to 0.2% and -1.7% to 6.0%, respectively. Any implementation or marketing costs would reduce both, and would likely make à la carte welfare-decreasing.

*We would like to thank Dan Ackerberg, Derek Baine, Lanier Benkard, Liran Einav, Catherine de Fontenay, Phillip Leslie, Michael Ostrovsky, Alessandro Pavan, David Pearce, Amil Petrin, Peter Reiss, Paulo Somaini, Alan Sorensen, Steve Stern, John Thanassoulis, Tracy Waldon, and seminar participants at numerous universities and conferences. We especially thank Yurukoglu's dissertation advisers Ariel Pakes, Luis Cabral, John Asker, and Allan Collard-Wexler. Robin Lee's collaboration on another project helped us to improve several aspects of the paper.

†Department of Economics, University of Warwick and Centre for Economic Policy Research (CEPR)

‡Graduate School of Business, Stanford University

1 Introduction

Bundling is widespread in multichannel television markets.[1] In theory, bundling can be a profitable form of price discrimination. It makes consumer tastes more homogenous and can facilitate surplus extraction, but has ambiguous effects on total welfare (Stigler (1963), Adams and Yellen (1976)). Regulations mandating *à la carte* pricing would radically alter the choice sets of the roughly 110 million U.S. television households who collectively spend more than $50 billion annually and watch an average of more than seven hours of television per day. This paper predicts the impact of such a regulation on the distribution of consumer and producer welfare.

There are widely differing opinions among policy makers, consumers, and industry participants about the effects of mandating à la carte pricing in the U.S.[2] This lack of consensus is partly because regulations mandating unbundling have not been implemented in enough similar circumstances to provide direct evidence.[3] Local experimentation would be informative about changes at the retail level, but à la carte would also affect industry-wide negotiations between content providers and distributors. We specify and estimate an industry model to evaluate à la carte pricing.

We model viewership, demand, pricing, bundling, and input market bargaining of multichannel television services. We first combine television viewership (ratings) data with bundle market shares and prices to estimate the distribution of household preferences for each of fifty cable television channels. We next estimate the input costs that distributors, such as Comcast or DirecTV, currently pay to content conglomerates, such as ABC Disney (which owns ESPN and The Disney Channel, among others) or Viacom (which owns MTV and Comedy Central, among others), for each of these channels using aggregate cost data and observed pricing and bundling decisions. The central innovation of our model is accounting for the change in distributors' input costs that result from bargaining between content and distribution in an à la carte world. To do so, we use the demand and cost estimates to estimate the parameters of a bilateral oligopoly bargaining model of the input market. Holding the estimated demand and bargaining parameters fixed, we simulate a world where distributors are forced to unbundle channels, critically allowing for the renegotiation of contracts between channel conglomerates and distributors.

In these counterfactual simulations, equilibrium input costs are an estimated 103.0% higher than when dis-

[1] Multichannel television refers to subscription-based television services. In the U.S., these are provided by cable television systems, direct-broadcast satellite (DBS) providers, and wireline video operators (especially incumbent telephone service providers). They are together called multichannel video program distributors (MVPDs).

[2] In addition to numerous articles in the popular press (e.g. Reuters (2003), Shatz (2006)), the Federal Communications Commission (FCC) has published two reports analyzing à la carte pricing (FCC (2004), FCC (2006)). The National Cable and Telecommunications Association (NCTA) has a webpage summarizing industry opposition to à la carte at http://www.ncta.com/IssueBrief.aspx?contentId=15. Supporters of à la carte include Consumers Union (http://www.consumersunion.org/pub/core_telecom_and_utilities/000925.html) and The Parents Television Council (http://www.howcableshouldbe.com/). According to a 2007 poll by Zogby, 52% of cable subscribers sampled supported à la carte pricing (http://www.zogby.com/news/readnews.cfm?ID=1377).

[3] Internationally, Canada, Hong Kong, and India have introduced various forms of regulations mandating unbundling in multichannel television markets, but idiosyncratic features of these regulations limit generalizations.

tributors sell bundles. These higher costs are passed into prices, offsetting the welfare benefits to consumers from being able to purchase individual channels. We estimate that, accounting for higher equilibrium input costs, consumer welfare changes between -5.4% and 0.2% and total welfare changes between -1.7% and 6.0%. Any implementation or marketing costs associated with à la carte would reduce both and would likely make à la carte welfare-decreasing in the short run.

The model has three types of agents: consumers, downstream distributors, and upstream channels. We estimate consumer preferences using both individual-level and market-level data on viewership, i.e. which channels consumers watch and for how long, and market-level data on bundle purchases, i.e. which bundles of channels consumers purchase, what channels they contain, and what prices are charged. We assume that the more a consumer watches a television channel, the more she is willing to pay for it. The viewership data provides the empirical evidence necessary for flexibly estimating a high-dimensional distribution of preferences for channels. The bundle purchase data provides the empirical evidence necessary to estimate how households trade off their utility from viewing channels with the price they have to pay for a bundle of those channels.

On the supply side, downstream distributors compete with each other by choosing both bundles and prices and by negotiating input costs with upstream channel conglomerates. We assume that observed prices and bundles are a Nash equilibrium given estimated preferences. We estimate input costs as those which make the Nash equilibrium assumption hold. We use the procedure in Pakes, Porter, Ho and Ishii (2007) to incorporate a subset of the necessary conditions implied by a 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.

To model the determination of input costs, we fix an industry bargaining protocol based on the model of Horn and Wolinsky (1988). The bargaining protocol features bilateral meetings between conglomerates of channels and distributors whose outcomes impose externalities on other firms due to downstream competition. We employ the equilibrium concept of contract equilibrium, as in Cremer and Riordan (1987), which requires that no pair of distributor and conglomerate would like to change their agreement given all other agreements. One notable empirical paper that also studies bargaining with externalities due to downstream competition is Ho (2009)'s analysis of hospital-HMO negotiations in the U.S. Our paper contributes to this line of research by using a bargaining model that includes Ho's take-it-or-leave-it offers as a special case. We estimate channel conglomerate-distributor specific bargaining parameters that produce the estimated input costs in equilibrium.

The estimated distribution of channel preferences replicates many features of the ratings data. For example, willingness-to-pay (WTP) for Black Entertainment Television (BET) is estimated to be higher on average for black households. Similarly, WTP for Nickelodeon and Disney Channel are estimated to be higher on average for family households. Average estimated own-price elasticities for basic cable, expanded basic cable, and satellite services are -4.1, -6.3, and -5.4, respectively.

Median estimated price-cost margins are 44%. We estimate that large distributors, such as Comcast, have about 17% lower input costs than small, independent distributors.

The estimated bargaining parameters reject take-it-or-leave-it offers as a model of the input market. On average, we estimate that distributors have higher bargaining parameters than channel conglomerates for small channel conglomerates, but that the situation is reversed for large channel conglomerates. Among distributors, small cable operators and satellite providers have slightly less estimated bargaining power than large cable operators.

We use these estimates to simulate the welfare effects of an à la carte pricing regulation. In the counterfactual simulation, we consider an economic environment with one large and one small cable market (each served by a single cable system), where the cable system and each of two "national" satellite distributors compete by charging a fixed fee and separate prices for each of the almost fifty cable television channels in our specification. We also simulate the welfare effects of theme tiers and a bundle-size-pricing regulation as in Chu, Leslie and Sorensen (2010). In all cases, we allow for input market renegotiation between channel conglomerates and distributors.

There are two countervailing forces that largely determine our results. First, for fixed input costs, unbundling unlocks consumer surplus. If we do not allow for input market renegotiation (i.e. input costs in an à la carte world stay at their bundle levels), forcing channels to be offered à la carte increases consumer welfare by an estimated 19.2% and reduces industry profits by 12.7%. Allowing renegotiation, however, increases costs by an estimated 103.0%. Prices follow suit, making the average consumer indifferent (increasing consumer surplus by 0.2%), increasing industry profits by 4.8%, and decreasing estimated total surplus from 4.1% to 2.4%.[4] Any implementation or marketing costs would reduce all of these and would likely make à la carte welfare-decreasing in the short run.

These estimates of the consequences of à la carte are for a baseline set of assumptions about demand, cost, and the nature of bargaining between channels and distributors that are described in detail throughout the paper. Where practicable, we have assessed the robustness of our conclusions to changes in these assumptions. For example, changes in assumptions regarding distributor markups under à la carte and the shape of and correlation between household preferences for channels yield qualitatively similar results: estimated consumer surplus changes between -5.4% and 0.2%, industry profits between 2.4% and 12.8%, and total surplus between -1.7% and 6.0%.[5]

Some of our assumptions cannot easily be evaluated, however. Perhaps the most important is that we infer greater utility for channels when they are watched more. We conduct monte carlo simulations in a simplified environment for data generated by an alternative viewership model that allows for channels which are viewed

[4] Bundle-size pricing and theme tiers are even worse for consumers (reducing welfare by an estimated 8.8% and 22.0%, respectively) as they still induce higher input costs, but do not permit households to select only the channels that they want.

[5] Bargaining outcomes are much more important for predicting surplus: if renegotiated input costs were to rise by half (double) the 103.0% we estimate, estimated consumer surplus would increase by 18.5% (fall by 27.6%). This merely emphasizes the importance of estimating a bargaining game and simulating counterfactual bargaining outcomes in order to accurately understand the effects of unbundling in television markets.

for a short time to have higher valuations than channels that are viewed for a longer time.[6] We find that our model predicts poorly outcomes for individual channels in this case, but still predicts well the overall (i.e. across-channel) welfare effects of à la carte. Another important assumption is that we analyze short-run effects taking the identities and qualities of channels as given. In the long-run, channels could enter, exit, and change how much they spend on programming, with important welfare effects in their own right. Finally, changes in consumer learning, preference formation, and/or so-called "behavioral effects" (e.g. Bertini and Wathieu (2008)) could also be important in a move from bundles to à la carte sales. As our data are not rich enough to evaluate these issues, we keep them as maintained assumptions and leave generalizations to future work. The interpretation of our results, however, should bear these assumptions in mind.

Related Work This paper is related to a number of empirical papers evaluating policy issues in these markets (Crawford (2000), Chipty (2001), Goolsbee and Petrin (2004)) as well as several papers addressing the identical topic. Rennhoff and Serfes (2009) develop a two-channel, two-distributor model with consumer preferences distributed uniformly on a circle to analytically study bundling and the wholesale market. Rennhoff and Serfes (2008) estimate a logit demand system for channels. In both studies, they conclude that à la carte regulations would likely increase consumer surplus, but the underlying modeling and distributional assumptions are too strong to adequately evaluate those claims. Crawford (2008) tests the implications of bundling in cable markets using reduced-form techniques. While suggestive, he does not identify the structure of channel demand required to estimate the welfare effects of bundling. The closest related work is due to Byzalov (2010). He estimates a model of demand for multichannel television using household-level survey data from a cross-section of four large DMA's in 2004. He finds that forcing cable distributors to offer theme tiers would decrease average consumer welfare at fixed wholesale prices. His household data are advantageous compared to our individual data in that they record the viewing behavior of all the adult television viewers in the household, but his market data are limited to a small sample of markets in 2004 rather than multiple thousands of systems over ten years as in this study.[7] Furthermore, he neither evaluates the welfare of full à la carte (i.e. having each channel itself available for sale) nor computes renegotiated input costs in his counterfactual analysis.[8]

2 Intuition for Results

The contribution of this paper can be understood by appreciating the insights of, and interaction between, two theoretical literatures in economics. The first evaluates the welfare consequences of bundling when input

[6]Channels offering sports programming, for example, may be watched less but valued more.

[7]Having observations on the adults within a household allows him to address the extent to which *within-household* correlation in tastes is an important for the discriminatory incentives to bundle.

[8]The results we present here are also related to results we have previously disseminated in working paper versions of this paper and related work. As our qualitative conclusions about the welfare effects of à la carte have changed in the process of conducting this research, we will describe how and why our conclusions have changed, but do so after introducing the ideas in the next section.

costs to the bundling firm are fixed (Stigler (1963), Adams and Yellen (1976)). The second models how those input costs are determined in a bilateral bargaining setting under oligopoly (Horn and Wolinsky (1988)). The ultimate welfare effects of à la carte depend on the interaction of the effects analyzed in these literatures, in particular on the magnitude of input cost increases that are likely to arise under à la carte. The three figures we now describe provide intuition for our results.

Figure 1 demonstrates the price discrimination incentive for bundling by a monopolist. Consider two goods with dispersed valuations and fixed marginal costs of zero given by the dashed lines in the figure. No matter the prices it charges, pricing each good individually requires a seller to miss out on the surplus from high valuation consumers willing to pay more than its price and low valuation consumers willing to pay less than its price but more than its cost. Compare that to the demand curve for the bundle. As long as valuations between the two goods are not perfectly correlated, consumers' valuation of the bundle will be less dispersed than those for the components, allowing the seller to capture more of the combined surplus with a single price. While we choose valuations that are highly negatively correlated in the figure to emphasize this point, it is quite general: à la carte regulations can unlock surplus and improve consumer welfare, for given input costs.[9]

The complication is that marginal costs can change under à la carte. Forgetting bundling for a moment, consider the determination of input costs for a single good in a bilateral monopoly with linear fee contracts, as in the two left-most panels of Figure 2. For a given input cost from the y-axis in the first panel, the downstream distributor in the second 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 bargaining literature cited above argues equilibrium input costs with linear fee contracts are determined as a function of a weighted geometric average of these two profits called the Nash product. The left panel traces out the Nash product for each possible input cost.[10] The equilibrium input cost maximizes the Nash product.

The third and fourth panels of Figure 2 combines the insights of these two literatures to determine input costs under bundling versus à la carte. It repeats the first two panels for two goods which have the same underlying mean valuations, but different dispersions. One can see that the equilibrium input cost for the more dispersed (à la carte) good is higher than that for the less dispersed (bundled) good. For many distributions of preferences, this drives up costs.[11]

[9]There is a long literature that has established this point for monopolists facing particular distributions of demand and cost (Adams and Yellen (1976), Schmalensee (1984)). Fang and Norman (2006) show that if preferences are symmetric and log-concave and average willingness-to-pay is greater than cost, then bundling is always more profitable than component sales.

[10]In this demonstration, we use equal weights. In our results, we estimate ζ_{fK}, the weighting for each pair of distributor and channel conglomerate.

[11]There is an additional, opposite effect of à la carte pricing on input costs. Bundling creates a negative externality in a channel's bargaining problem as a higher input cost weakens demand for the other channels in the bundle. This externality makes input costs higher under bundles; eliminating it pushes input costs *lower* under à la carte. On average, we find input costs rise considerably, so in aggregate this externality effect is dominated by the niche pricing effect described in the text. However, for some channels it is the dominant effect.

The key to understanding the welfare effects of à la carte is to know how much input costs would rise under mandatory à la carte. If modest, the insights of the bundling literature likely obtain and à la carte could be consumer and total welfare-enhancing. If extreme, prices under à la carte will also be high, making it much more likely to be welfare-reducing. How much input costs rise under à la carte in practice particularly depends on the structure of preferences for individual channels and the relative bargaining power of channels and distributors. These are the focus of our econometric estimation in the sections to follow.[12]

3 The Data

We divide our data into two categories: market data, which measure households' purchasing decisions or firms' production decisions, and viewership data, also called ratings, which measure households' utilization of the cable channels available to them.

Our market data comes from two sources: Warren Communications and SNL Kagan. Warren produces the Television and Cable Factbook Electronic Edition monthly (henceforth Factbook). The Factbook provides data at the local cable market level on the composition of cable television bundles, their prices and market shares, cable system ownership, and other system characteristics. SNL Kagan produces the Economics of Basic Cable Networks yearly (henceforth EBCN). EBCN provides data at the level of channels on a variety of revenue, cost, and subscriber quantities.

Cable System (Factbook) and Satellite Data Our Factbook sample spans the time period 1997-2007. The Factbook collects its data by telephone and mail survey of cable systems. The key data in the Factbook are the cable system's bundle compositions, the prices of its bundles, the number of monthly subscribers per bundle, the number of homes passed by the cable system, and the ownership of the system.

Table 1 and part of Table 2 provide summary statistics for the Factbook data. An observation is a system-bundle-year, e.g. NY0108's Expanded Basic in 2000. We observe over 25,000 system-bundle-years, based on over 19,000 system-years from over 8,000 systems. Most systems in our data offer a single bundle, while the majority of the rest offer just two bundles. Much of our data comes from early in the sample period when fewer offerings were the norm.

[12] The trade-off between unbundling all offered TV channels (i.e. Full à la carte, or Full ALC) and higher input costs due to re-negotiated bargaining under à la carte is the driving force in predicting consumer welfare benefits of à la carte. This paper is the combination of what was two separate research papers, each looking at measuring the welfare benefits of à la carte. The first paper, last circulated in February 2009, by both authors, allowed Full ALC, but not input bargaining effects and, like previous work by the first author using similar assumptions (e.g. Crawford (2008)), unsurprisingly found significant consumer welfare benefits. The second paper, last circulated in April 2009, by the second author (Yurukoglu (2009)), introduced the bargaining model and input bargaining effects, but couldn't do so while allowing Full ALC, focusing instead on a blend of Bundle-Sized Pricing (Chu et al. (2010)) and a few channels being offered ALC whose effects were similar to pure bundling. This paper unsurprisingly found very modest consumer welfare benefits. It is only in the current paper (combining those research projects) that we have developed methods to flexibly allow both Full ALC and input bargaining effects to permit the data to tell us the relative importance of each.

For each of these bundles and by market type, Table 1 reports the average price of the bundle in 2000 dollars, it's market share, and the number of cable channels offered. In markets with two or more bundles, the average Basic service in our data costs about $13.50 and offers about 9 cable channels and the average Expanded Basic bundle costs around $30.00 and offers about 30 cable channels.[13]

There is variation in the composition of bundles across markets and over time. Table 2 presents the share of systems in our sample that offer each of the channels in our specification. The first column indicates whether the channel is carried on any tier of service, while the second column indicates whether the channel is offered on the basic tier. For example, ESPN is carried by almost all systems (96.7%) in our data. Of these, most (76.7%) carry it on Basic Service. Smaller channels are frequently offered on Digital Service.

Unlike for cable service, satellite offerings do not vary by geography. We collected satellite menus and prices by hand. We then matched this to aggregate satellite market share data at the DMA level from Nielsen Media Research.[14]

Aggregate Channel (SNL Kagan) Data We use the 2006 edition of the Economics of Basic Cable Networks (EBCN). The 2006 sample covers 120 cable channels with yearly observations dating back to 1994 when applicable. Information collected includes total subscribers, license fee revenue, advertising revenue, and ownership. The data are collected by survey, private communication, consulting information, and some estimation. The exact methods used are not disclosed. The key variables we use are the average input cost (denoted τ_c for a given channel c later in the paper), and the advertising revenue for each channel. The average input cost for a channel is its license fee revenue divided by the number of subscribers. It measures how much distributors are paying for the channel per subscriber, averaged across distributors. In 2007, this ranged from $3.26 for ESPN to $0.03 for MTV2 for the roughly fifty channels in our model.

Viewership Data Our viewership data comes from two sources: Nielsen and Mediamark. The Nielsen data is DMA-level tuning (viewing) data. The Mediamark data is individual-level survey data.

Nielsen DMA Tuning Data The Nielsen data comes from the 56 largest DMA's for about 50 of the biggest cable channels over the period 2000-2006 in each of the "sweeps" months of February, May, July, and November. The main variables are the DMA, the program, the channel, and the program's rating.. The rating is the percentage of households with at least one television in the DMA viewing the programming on that channel.

We aggregate the information across programs on each channel within each month of our data. Thus an

[13] Digital basic packages were made possible by cable systems investments in digital infrastructure in the late 1990's and 2000's. This dramatically 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, often called digital basic.

[14] Designated Market Areas, or DMAs, correspond to local broadcast television coverage areas. There are usually several cable systems within a DMA.

observation is a channel-DMA-year-month, e.g. the average rating for ESPN in the Boston DMA in February, 2004. We have 1,482 such combinations. The third column in Table 2 presents the average rating for each of the channels in our analysis.

We observe that channels' ratings vary from DMA to DMA and within DMA across months and years. One important type of variation we use is how ratings vary with the demographic composition of a DMA. We focus on six demographic factors: Family status, Income, Race, Education, and Age.[15] Figure 6 in Appendix B.2 provides an illustrative example of the impact demographic characteristics can have on ratings by comparing average ratings for Black Entertainment Television (BET) across markets. Table 11 in Appendix B.2 reports correlations in the DMA-month-year ratings across a subset of cable channel pairs. Correlations in viewing from our household-level data show similar patterns.

Mediamark Individual level Data The Mediamark data comes from surveying a random sample of consumers in the US about their media usage, consumer behavior, and demographics. They survey roughly 25,000 individuals per year. Our data spans the years 2000 to 2007. Individuals report how many hours they watch each of over 75 cable channels in a given week.

In columns four and five of Table 2, we present the mean and the standard deviation of the fraction of households reporting viewing a certain channel per hour.[16] This is analogous to an average Nielsen rating for that channel and for that reason we call them "ratings" in the table. The final column reports what fraction of households report positive viewing of each channel. In industry parlance, this is known as the "cume," short for cumulative audience.

Data Quality Issues About four-fifths of the possible observations in the Factbook on market share and price for cable bundles are either missing, not updated from the previous year, or both.[17] We assume this data is missing at random conditional on the observable characteristics of the system. Most systems show up at least once in the time period of the data set.

We only observe the aggregate satellite market share at the DMA level. For the demand estimation, we assume that there is only one satellite firm offering DirecTV's Total Choice package. In reality, both DirecTV and Dish offer three to four tiers of service each.

The Mediamark data is at the individual level while our model is at the household level. To use this data to estimate our model, we create synthetic households by matching individuals to households based on observable characteristics like age, cable or satellite subscription, marital status, household income, and race.[18] For

[15] We follow U.S. Census definitions for each of these variables. Table 12 in Appendix B.2 reports sample statistics across the 56 DMAs for which we have ratings data.

[16] These are fictional households are created from the real individual data as detailed in the Data Quality section immediately below.

[17] Appendix B.2 discusses data quality issues for the datasets used in this paper in more detail.

[18] This is one advantage of the data in Byzalov (2010): it reports the viewing for all adult members of a household, eliminating

each observation, we randomly draw an individual level observation. We then draw more individuals with similar characteristics to fill in the other members of the reported household size. If several individuals could fit into a given household, we choose at random. If individuals who share the same tastes in television tend to marry, then with this procedure we will overestimate the number of channels watched by households, while if opposites attract, we will underestimate that number.

4 The Industry Model

The industry model predicts household 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.

In **stage 1**, channels and distributors bargain bilaterally to decide input costs; in **stage 2** distributors set prices and bundles; in **stage 3** households make purchases; and in **stage 4**, households view television channels. We start from the last stage and work backwards.

4.1 Household Viewing

Let j index a bundle of programming being offered by cable system n in DMA d in month-year m (e.g. Comcast Digital Basic in Arlington, VA in the Washington, DC DMA in November 2003) and let b_{dnm} be the set of all such bundles.[19,20] We will suppress the market subscripts n, d, and m for the moment. Let c index channels and let C_j be the set of channels offered in bundle j. We assume the utility to household i from spending their time watching television and doing non-television activities has the Cobb-Douglas in logs form:

$$v_{ij}(t_{ij}) = \sum_{c \in C_j} \gamma_{ic} \log(1 + t_{ijc}) \tag{1}$$

where t_{ij} is a vector with components t_{ijc} which denote the number of hours household i watches channel c when the channels in bundle j are available, and γ_{ic} is a parameter representing i's tastes for channel c.[21]

the need for this kind of imputation.

[19]For convenience, we index month-year combinations (e.g. November, 2003; May, 2004; November, 2004) by the single index, m.

[20]We have two geographic identifiers: cable markets n and Nielsen DMAs d. This is necessary due to the different levels of geographic aggregation in our data.

[21]One could experiment with richer models of time allocation. For example, one could model a sequence of discrete choices of which channel to watch in every fifteen minute period. The combination of Nielsen ratings and recently developed set-top box tuning data would allow the researcher to estimate such a model. A richer model would allow us to test our viewership model against data which details time-of-day viewing. Additionally, it would allow one to transparently impose additional assumptions such as that viewing during prime time is more valuable than viewing during mid-morning. Unfortunately, our individual-level

We will later estimate the distribution of γ allowing for positive or negative correlations in tastes for pairs of channels. Households may opt to not watch any channel, and we call this state channel 0, $0 \in C_j\ \forall j$, with t_{ij0} the amount of time household i spends on non-television activities and γ_{i0} their preferences for such activities.

Each household i solves:

$$\begin{aligned} \max_{t_{ij}} \quad & \textstyle\sum_c \gamma_{ic} \log(1 + t_{ijc}) \\ \text{subject to} \quad & \textstyle\sum_c t_{ijc} \leq T \end{aligned} \tag{2}$$

with the additional restrictions that the time spent watching any channel must be non-negative, and the time spent on channels not in bundle j is zero.

The solution to this maximization problem yields household i's indirect utility from viewing the channels in bundle j:

$$v^*_{ij}(\gamma_i, C_j) \quad = \textstyle\sum_{c \in C_j} \gamma_{ic} \log(1 + t^*_{ijc}) \tag{3}$$

Discussion We infer how much a household values a channel relative to other channels based on how much time they spend watching that channel relative to other channels. This would not be good assumption, for example, if households valued the option of watching The Weather Channel in case of bad weather, but never watch under normal circumstances or if programming on some channels is highly valued but only watched for a short period of time relative to programming on other channels (e.g. high-profile sporting events).

Because channels are uniformly sold in large bundles, bundle data alone doesn't provide enough variation to separately estimate household demand for individual channels. Viewing data does provide channel-specific variation, but no prices. It is the combination of these types of data - and the assumption that viewing time informs value - that enables us to quantify the welfare benefits of à la carte policies.

To address the likely consequences of this assumption for our results, we conduct a monte carlo exercise as part of our robustness analysis in Section 7 that allows for channels watched a short time to be valued more than channels watched for longer periods. A brief summary of our findings is provided there and a full description of this exercise and its results is provided in Appendix B.2.

4.2 Bundle Purchases

A household's choice of cable bundle will depend on their utility from having access to the channels in that bundle, v^*_{ij}, as well as other characteristics of the bundle and cable system such as the bundle's price. We

viewership data does not contain time of day viewing. Because of this data limitation and the increased computational requirements for estimating the richer model, we employ the simple Cobb-Douglas model of time allocation presented in the text. In Appendix B.2, we explore the implications of a richer viewership model which allows for consumers to value channels they watch a short time more than channels they watch for longer periods.

assume the utility household i derives from subscribing to bundle j in market n in DMA d in month m as:

$$u_{ijndm} = v^*_{ijndm} + z'_{jndm}\psi + \alpha_i p_{jndm} + \xi_{jndm} + \epsilon_{ijndm} \quad (4)$$

where, $v^*_{ijndm} = v^*_{ijndm}(\gamma_i, C_{jndm})$, from (3), represents the indirect utility to household i from viewing the channels available on bundle j, p_{jndm} is the monthly subscription fee of bundle j, and z_{jndm} are other observed system and bundle characteristics of bundle j in market n, DMA d, and month m. For convenience, we will sometimes refer to this triple as "market ndm". $\alpha_i = \alpha + \pi_p y_i$, with y_i household i's income, is a taste parameter measuring the marginal utility of income. ψ is a parameter measuring tastes for system and other bundle characteristics. ξ_{jndm} and ϵ_{ijndm} are unobserved portions of household i's utility. We assume that the unobserved term has a component which is common to all households in the market, ξ_{jndm}, and an idiosyncratic term, ϵ_{ijndm}. We further assume that the idiosyncratic term is an i.i.d. draw from a type I Extreme Value distribution whose variance we set to 1.[22]

The components of z_{jndm} include by which MSO, if any, the bundle is being offered, the year the bundle is being offered, and bundle name dummies (e.g. "Basic", "Expanded Basic", etc.). ξ_{jndm} represents the deviation of unobserved demand shocks or bundle attributes from the MSO-year-bundle name mean. These unobserved attributes in our data include price and quality of tied Internet service, high definition (HD) service, promotional activity, technical service, and quality of equipment. Theory predicts that these unobservable attributes will be correlated with price. In the estimation section, we will use instrumental variables to disentangle the effect of price from any correlation with unobservable attributes.

Define $\delta_{jdnm} = z'_{jndm}\psi + \alpha p_{jndm} + \xi_{jndm}$ and $\mu_{ijndm} = v^*_{ijndm} + \pi_p y_i p_{jndm}$. Let F^n be the distribution of household preferences and demographics in market n. By the distributional shape assumption on ϵ_{ijndm}, the model's predicted market share for bundle j in market n in DMA d in month m is:

$$s_{jndm} = \int \frac{exp((\delta_{jndm} + \mu_{ijndm}))dF^n(i)}{1 + \sum_{k \in ndm} exp((\delta_{kndm} + \mu_{ikndm}))} \quad (5)$$

Our model assumes that the amount of time spent by households watching channels is informative for what they are willing to pay for access to those channels.[23] We also assume that all households have non-negative willingness to pay for channels.

4.3 Supply: Downstream Distributors

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

[22]The inclusion of viewing behavior embodied in v^*_{ijndm} has two implications for our bundle purchase model. First, we normalize the scale of utility by setting the parameter on v^* to 1. Second, it allows us to estimate the variance of ϵ (which is normally not feasible as that is chosen as the utility scale normalization). In practice, this estimated variance was small relative to the variance of the other elements of utility, so we (also) set it to one. We retain it in the model as it provides a useful computational role in the econometric estimation by smoothing demand as a function of the underlying parameters.

[23]We discuss the reasons for and consequences for our results of this important assumption at the end of this section.

The profit of a distributor before fixed costs is:

$$\Pi_{fndm}(\mathbf{b_{ndm}}, \mathbf{p_{ndm}}) \;=\; \sum_{j\in \mathbf{b_{fndm}}} (p_{jndm} - \sum_{c\in C_{jndm}} \tau_{fc}) s_{jndm}(\mathbf{b_{ndm}}, \mathbf{p_{ndm}}) \tag{6}$$

where f denotes distributor, n market, d DMA, m month, and j bundle. $\mathbf{b_{ndm}}$ is a list of offered bundles in market ndm with corresponding prices $\mathbf{p_{ndm}}$ and $\mathbf{b_{fndm}}$ are the bundles offered by firm f. τ_{fc} are distributor-channel specific license fees. Taking a distributor's perspective, we refer to these as "input costs" throughout this paper. Distributor f pays channel c a payment of τ_{fc} for every household which receives channel c from firm f. Following the nature of programming contracts in the industry, these vary by firm and channel, but not across the markets served by firm f.

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

Nash Assumption $\forall f$ and $\forall ndm$, $\mathbf{b_{fndm}}$ and $\mathbf{p_{fndm}}$ maximize $\Pi_{fndm}(\mathbf{b_{ndm}}, \mathbf{p_{ndm}})$ given $\mathbf{b_{-fndm}}$ and $\mathbf{p_{-fndm}}$.

The Nash assumption implies that bundle prices satisfy the downstream 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.

We 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 our estimated parameters. Multiple Nash equilibria would negatively affect both the estimation of bargaining parameters and the simulation analysis of unrealized policies. While we cannot prove uniqueness, we do numerically search for multiple equilibria by changing the starting values when computing an equilibrium by best-response dynamics and do not find multiple equilibria.

4.4 Supply: Bargaining Between Distributors and Channel Conglomerates

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. This paper's environment differs from those models because payoffs depend on outcomes of bilateral negotiations that firms are not party to. These cross-negotiation 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. de Fontenay and Gans (2007) extend these models to allow for arbitrary numbers of agents on both sides of the market.

We too model this situation as a game involving the upstream channels, or conglomerates of channels, and the downstream distributors. Distributors and conglomerates meet bilaterally. Following industry practice, we assume distributors (MSOs) negotiate on behalf of all their component systems and channel conglomerates bargain on behalf of their component channels. They bargain à la Nash 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.

We 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, payments are linear, but contain other provisions as well: descriptions of the service to be provided by each side, standards for technical service, marketing agreements, most favored nation clauses, division of advertising spots, tiering requirements, and auditing, confidentiality, and severability clauses. However, few contain fixed monetary transfers, and if they do, they are negligible with respect to the contract's total value. We model the contracts as only a linear fee for each distributor and channel.[24]

Let $\Psi = \{\tau_{fc}\}$ be a set of input costs, a scalar for each pair of distributor and channel. In the bargaining stage, each conglomerate of channels and distributor meets separately and simultaneously. We denote a conglomerate by K and a channel by c. Let τ_{fK} be the vector of input costs for conglomerate K. We assume these meetings result in the asymmetric Nash bargaining solution. In each bilateral meeting, τ_{fK} maximizes firm f and conglomerate K's bilateral Nash product:

$$NP_{fK}(\tau_{fK};\Psi_{-fK}) = \left[\Pi_f(\tau_{fK};\Psi_{-fK}) - \Pi_f(\infty;\Psi_{-fK})\right]^{\zeta_{fK}} \left[\Pi_K(\tau_{fK};\Psi_{-fK}) - \Pi_K(\infty;\Psi_{-fK})\right]^{1-\zeta_{fK}} \quad (7)$$

where Π_f is the sum over markets (ndm) of firm f's profit function in (6) and

$$\Pi_K(\tau_{fK};\Psi_{-fK}) = \sum_{c\in K}\left(\sum_f \tau_{fc}Q_{fc}(\Psi)\right) + r_c^{ad}t_c(\Psi)$$

is conglomerate K's profit function before fixed costs. $Q_{fc}(\Psi)$ is the total number of subscribers of channel c coming from distributor f and r_c^{ad} is the advertising revenue of channel c per household hour watched. The endogenous viewership, $t_c(\Psi)$, is recomputed in every downstream equilibrium using the consumer demand and viewership model. In words, the conglomerate profit function is the sum over distributors of license fee plus advertising revenue. Advertising revenue depends on the advertising rates and endogenous viewership of the conglomerate's channels. If there is no agreement between a distributor and a conglomerate, then the input cost for each channel in the conglomerate is positive infinity.

Negotiations are simultaneous and separate, so Ψ_{-fK}, the set of all other input costs, is not known but conjectured. ζ_{fK} is the bargaining parameter of distributor f when meeting conglomerate K. Allowing

[24] 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 such competition.

$\zeta_{fK} \neq 0.5$ distinguishes asymmetric from symmetric Nash bargaining. Setting ζ_{fK} to zero is equivalent to assuming Nash-Bertrand pricing behavior by the upstream firms.

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

The interpretation of this equilibrium, due to Horn and Wolinsky (1988), is a Nash equilibrium between Nash bargains. To paraphrase, consider a simultaneous move game where the players are the bargaining pairs, each pair's strategy is τ_{fK}, and each pair's payoff is its Nash product. The bargaining equilibrium is the Nash equilibrium of that game. This setup does not allow for advantages due to informational asymmetries. Each distributor and each conglomerate sends separate representatives to each meeting. Once negotiations start, representatives of the same firm do not coordinate with each other.[25] We view this absence of informational asymmetries as a weakness of the bargaining model. However, in return we gain tractability in determining how the threat of unilateral disagreement determines input costs in a bilaterally oligopolistic setting.

Another 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, we assume that agreements are binding in all contingencies. In previous versions of this paper, we have solved alternative cases where if a pair disagrees, all other firms renegotiate conditional on the disagreeing pair dropping out forever. This case is reminiscent of the reasoning in the Shapley value.[26] This alternative model generated different estimates of bargaining parameters, but did not affect our ultimate results. Solving this alternative game is computationally more challenging because one must compute payoffs for every possible configuration of agreement or disagreement. Without more industry specific information on what might happen to other negotiations when a pair disagrees, and given that both models deliver similar ultimate conclusions, we chose the simpler model.

In our baseline specification, we treat each conglomerate as an indivisible block of channels. This implies, for example, that if bargaining breaks down between ABC Disney, which owns ESPN, ESPN 2, Disney Channel, ABC Family, SOAPNet, and other channels, and Comcast, then Comcast will not carry any of the ABC Disney channels. We also have solved a specification where we treat each channel as an individual firm. We assume that the disagreement profits 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. Recent details of negotiations which became public provide evidence for both assumptions: Viacom threatened to pull all of its channels, including MTV, Comedy Central, and Nickelodeon, during negotiations with Time Warner Cable in late 2008, whereas Comcast's content division pulled Versus from DirecTV in 2009 following an unsuccessful negotiation, but continued to serve its other channels, such as Golf Channel and E!, through DirecTV. How multi-product firms decide between potentially complex bargaining threats is an open question.

[25] As a separate issue, we also ignore moral hazard. For example, we ignore the imperfectly observable choice of effort exerted by channels to make compelling programming following an agreement. Descriptions of the programming are often written into the agreements, but it is not clear if there is a conflict between the two parties about these terms. Linear fees also may help resolve any more hazard issues upstream.

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

5 Estimation

We first estimate the distribution of preferences for channels, γ_i, using ratings data, jointly with the distribution of marginal utility of income, α_i, and non-price preference parameters, ψ, using market share, price, and bundle characteristics data. We then use these demand estimates to separately estimate a parameterized cost function which predicts an input cost, τ_{fK}, for each pair of distributor f and channel conglomerate K. Finally, given the estimated demand and cost parameters, we choose bargaining parameters, ζ_{fc}, for each pair so that the bargaining model induces the estimated set of input costs in equilibrium. While it would be efficient to estimate all the parameters jointly, we found it simpler to code and estimate the model as this sequence of separate steps.

5.1 Household Preference Parameters

We jointly estimate a parameterized distribution of γ with a parameterized distribution of α_i and non-price preference parameters, ψ. The moments used in estimation are: (1) the fraction of households that watch zero hours by channel for the eight combinations of three demographic groups (black, age, and family), (2) mean hours watched per household per channel by demographic group, (3) the covariance in DMA ratings with DMA mean demographics, (4) mean hours watched per household per channel, (5) the cross channel covariance in household hours watched, (6) the aggregate cable and satellite market share by income level, and (7) the covariance of demand-side instruments, Z_{jndm} with the unobserved demand shock ξ_{jndm}.

Household i's time spent viewing the programming on bundle j, t_{ijndm} depends on their vector of channel preferences, γ_i, and the channels available on bundle j, C_{jndm}. The ratings data are measurements of time spent viewing at the individual and market level. We estimate the distribution of γ by matching moments of the model's predictions of time spent viewing to moments of the ratings data. Relative to the existing literature on empirical demand estimation, we choose a novel structure of household preferences for channels, γ. We parameterize the distribution of γ as:

$$\gamma_i = \chi_i \circ (\Pi o_i + v_i)$$

where χ_i is a vector whose components are indicator random variables

$$\chi_{ic} = \begin{cases} 0, & \text{w. prob } \rho_{o_i c} \\ 1, & \text{w. prob } 1 - \rho_{o_i c} \end{cases}$$

In words, each household's vector of channel preferences consists of individual channel preferences, γ_{ic}, which is zero for a given channel with some probability depending on household demographics. If γ_{ic} is not zero, it is a random variable whose mean depends linearly on household demographics Πo_i, where o_i is a vector of demographic attributes of household i. There is a layer of unobservable heterogeneity in channel preferences due to the vector v_i which we assume is drawn from a multidimensional distribution named

G with exponential marginal distributions (whose parameters Λ we estimate) and a correlation structure described by a correlation structure Σ (which we also estimate). With this parametrization, the household maximization in Equation (2) yields $\hat{t}_{ijcndm}(\Pi, \rho, \Lambda, \Sigma)$, each household's time watched of channel c in bundle j.

This specification of tastes for channels captures the idea that some people simply don't value some channels. This happens with probability $\rho_{o_i c}$. For those that do, we assume preferences are distributed as an exponential distribution. Figure 3 demonstrates that viewing for the news channel CNN in our individual-level data is consistent with these assumptions. Similar patterns arise for all the channels in our analysis.

One can only observe ratings data for channels which a household has elected to receive. This introduces a selection issue: we are likely to observe the viewing decisions of those households with strongest tastes for channels. We accommodate this "selection into bundles" by matching moments of the model's predictions of time spent viewing conditional on bundle choice to ratings data which exhibit the same conditioning. The conditioning on bundle choice requires knowing parameters from the model of bundle choice (stage three of our model, given in equation (4)). We jointly estimate the parameters of the distribution of channel preferences together with bundle choice parameters as in Lee (2010). This allows us to recover the unconditional distribution of preferences for channels, an important element for our counterfactual simulations.

Identification The population moments of the model's predicted time spent viewing are sensitive to a limited set of parameters. One may casually think of those moments' observed counterparts as "empirically identifying" these parameters. Using this terminology, $\rho_{d_i c}$ is empirically identified by (1), the fraction of households that watch zero hours by channel by demographic group, Π by (2), the mean hours watched by household by demographic group, and (3), the covariance in DMA ratings with DMA demographics, G's marginal distribution exponential parameters by (4), the mean and variance in hours watched by household, and the correlation structure of G by (5), the cross channel covariance of household hours watched (net of variance attributed to demographics). Identification of the other demand parameters is discussed below.

Positive correlation for a pair of channels could arise if a certain demographic group watches both channels, or even in the absence of demographic patterns, if those who watch one of the channels also watch the other. Negative correlation could arise if exclusive demographic groups watch each channel, for example if rich households watch one of the channels and poor households the other, or even in the absence of demographic patterns, if those who watch one channel don't watch the other.

We parameterize the distribution of α_i as $\alpha_i = \alpha + \pi_p y_i$ where y_i is household i's income. We estimate α, π_p, and ψ as in Berry, Levinsohn and Pakes (2004) and Petrin (2003). This part of the estimation is based on Equation (5). For given values of π_p and the distribution of γ, we find the values of δ_{jndm} which equate observed market shares with predicted market shares using the contraction mapping from Berry, Levinsohn and Pakes (1995). Given δ_{jndm}, we estimate α and ψ by linear instrumental variables regression using instrument vector, $Z_{jndm} = [z_{jndm} \; w_{ndm}]$.

We assume observed non-price product characteristics (dummy variables for non-channel bundle characteristics such as firm, year, and tier name), z_{jndm}, are independent of ξ_{jndm}. We accommodate the endogeneity of price by instrumenting for it with w_{ndm}, where w_{ndm} is the average price of other cable systems' bundles within the same DMA as cable system n. Following Hausman (1996), these are often called "Hausman" instruments. These instruments have been used for demand estimation in settings such as Hausman, Leonard and Zona (1994) and Nevo (2001). They will be valid instrumental variables if, for bundle j in market n, the two standard conditions hold. First, they need be correlated with the price of bundle j in market n. This will be true if marginal costs are correlated with prices within n's DMA outside market n. Labor costs and advertising rates are cost shifters that are plausibly correlated within DMAs, suggesting this is likely to be satisfied. Second, they need be uncorrelated with the unobserved demand shock in market n, ξ_{jndm}. As discussed in Section 4.2 above, we anticipate ξ to contain unobserved characteristics of that system's types and quality of service (e.g. Internet access). Cable systems are physically distinct entities for which local managers have wide authority, so bundle prices should be uncorrelated with non-competing bundles' unobservable characteristics. Of course, other instruments are possible; we consider and evaluate several in Section 6. π_p is empirically identified by the total cable and satellite market share by income level.

The model's predicted time spent by household i watching channel c when subscribing to bundle j is given by $\hat{t}_{ijcndm}(\delta, \pi_p, \Pi, \rho, \Lambda, \Sigma)$ and depends on the data in addition to the indicated dependence on model parameters. The model's predicted market share for household i for bundle j is $\hat{s}_{ijndm}(\delta, \pi_p, \Pi, \rho, \Lambda, \Sigma)$. Explicitly, the moment conditions used in estimation are:

$$\begin{array}{l} (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \\ (7) \end{array} \left[\begin{array}{l} \frac{1}{N_{ndm}} \sum_{ndm} \frac{1}{N_{ondm}} \sum_{i=1}^{N_{ondm}} \left(\sum_{j \in b_{ndm}} \mathbf{1}_{\{\hat{t}_{ijcndm} > 0\}} \hat{s}_{ijndm} \right) - r_{co}^{\text{cume}} \\ \frac{1}{N_{ndm}} \sum_{ndm} \frac{1}{N_{ondm}} \sum_{i=1}^{N_{ondm}} \left(\sum_{j \in b_{ndm}} \hat{t}_{ijcndm} \hat{s}_{ijndm} \right) - t_{co} \\ \frac{1}{D} \sum_{d=1}^{D} (\hat{t}_{cd} - \bar{t}_c)(o_d - \bar{o}) - \sigma_{r_{cd}, o_d} \\ \frac{1}{D} \sum_{d=1}^{D} \hat{t}_{cd} - r_{cd} \\ \frac{1}{N_{ndm}} \sum_{ndm} \frac{1}{N} \sum_{i=1}^{N} \left(\sum_{j \in b_{ndm}} (\hat{t}_{ijcndm} \hat{s}_{ijndm} - \bar{t}_c)(\hat{t}_{ijc'ndm} \hat{s}_{ijndm} - \bar{t}_{c'}) \right) - \sigma_{t_c, t_{c'}} \\ \frac{1}{N_{ndm}} \sum_{ndm} \sum_{j \in b_{ndm}} \frac{1}{N_{ondm}} \sum_{i=1}^{N_{ondm}} \hat{s}_{ijndm} - s_o \\ \frac{1}{N_{ndm}} \sum_{ndm} \sum_{j \in b_{ndm}} \xi_{jndm} Z_{rjndm} \end{array} \right] = 0$$

where $\sum_{ndm}$ is the sum over markets, DMAs, and months in our data, N_{ndm} is the number of such market-DMA-months, $\hat{t}_{cd} = \frac{1}{N_{nm}} \sum_{nm} \sum_{j \in b_{ndm}} \frac{1}{N} \sum_{i=1}^{N} \hat{t}_{ijcndm} \hat{s}_{ijndm}$ is the average time spent watching channel c in DMA d and $o_d = \frac{1}{N_{nm}} \sum_{nm} \sum_{j \in b_{ndm}} \frac{1}{N} \sum_{i=1}^{N} o_{indm}$ is the average of demographic o in DMA d in the third moment (with $\bar{t}_c$ and $\bar{o}$ the across-DMA averages of those), Z_{rjndm} is the r^{th} instrument in Z_{jndm}, and we've suppressed the dependence of predicted time and market shares on the model's parameters and data to economize on space. On the right-hand side of the first six moment conditions are the corresponding moments in our data. r_{co}^{cume} is the share of MRI households of demographic o that have positive viewing to channel c, t_{co} is the average time MRI households of demographic o spend watching channel c, σ_{r_{cd}, o_d} is the across-DMA covariation in Nielsen ratings for channel c and demographic o, r_{cd} is the across-month average

Nielsen rating for channel c in DMA d, $\sigma_{t_c,t_{c'}}$ is the covariation in MRI households' time spent watching each pair of channels, c and c', and s_o is the market share for cable (and, separately, satellite) by demographic.

N_{ondm} is the total number of households who have demographic characteristic o in market ndm and D is the total number of DMA's. The set of demographic characteristics we use depends on the set of moments. For the set of moments associated with the first row, we use each of eight combinations of black, family, and whether the head of household is aged over 55. For the set of moments associated with the second and third rows, we use whether the household is a family or not, income level, race, whether the head of household has a bachelor's degree, and the age of the head of household. For the moments associated with the second-to-last row, we use income quartiles only. For convenience, the labeling of the moments to the left of the brackets corresponds to their description at the beginning of this subsection.

5.2 Cost Estimation

National-average input costs, the necessary conditions implied by Nash equilibrium in prices and bundles, and the observed prices and bundles identify input costs. National-average input costs are direct evidence. The rest is indirect evidence; what could input costs have been given the Nash assumption and observed prices and bundles?

We parameterize τ_{fc} as a function of channel characteristics scaled by a function of firm and channel characteristics:

$$\hat{\tau}_{fc}(\eta, \varphi) = (\eta_1 + \eta_2 \tau_c) exp(\varphi_1 MSOSIZE_f + \varphi_2 VI_{fc})$$

where τ_c is the (observed) Kagan average input cost for channel c, $MSOSIZE_f$ is firm f's total number of subscribers, and VI_{fc} is the ownership share firm f has in channel c.[27] While different channels may have different base rates, we assume the functional form of the effect of distributor size and vertical integration on input costs 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 with which it is not vertically integrated. This is a restrictive parametrization, even more so because we don't allow the coefficients to vary by year. It does however capture the distributor size effect which is the most important factor driving differences in distributor's fees for a given channel, and common knowledge in the industry.

A weighted average of τ_{fc} over firms predicts the national-average input cost for each channel c. The Kagan EBCN data set's channel input costs, τ_c, are the empirical counterpart of these averages. The first set of moment conditions is that the model's predicted aggregate input costs should equal observed aggregate input

[27] This information was collected from a number of different sources, primarily various years of SNL Kagan's EBCN and historical issues of *Multichannel News*.

costs with deviations from this relationship capturing measurement error in τ_c:[28]

$$E_f[\hat{\tau}_{fc}(\eta,\varphi)] - \tau_c = 0$$

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

$$\frac{d\Pi_{fndm}(\mathbf{b_{ndm}},\mathbf{p_{ndm}})}{dp_{kndm}} = \sum_{j\in B_{fndm}} (p_{jndm} - \sum_{c\in C_{jndm}} \tau_{fc}) \frac{ds_{jndm}(\mathbf{b_{ndm}},\mathbf{p_{ndm}})}{dp_{kndm}} + s_{kndm}(\mathbf{b_{ndm}},\mathbf{p_{ndm}})$$

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 a Nash equilibrium for each firm in each market, given all other bundles and prices. Given the estimated demand parameters and observed prices and bundles, we solve for the implied marginal cost of each bundle, $\sum_{c\in C_{jndm}} \tau_{fc}$, which we call $\hat{mc}_{jndm}$. The second set of moment conditions is that the difference between $\hat{mc}_{jndm}$ and $\sum_{c\in C_{jndm}} \hat{\tau}_{fc}(\eta,\varphi)$ should have zero covariance with the size of bundle j's MSO and the number of own vertically integrated channels included in bundle j and year dummy variables and tier name dummy variables. This is analogous to adding a bundle-specific error term measuring unobserved shocks to bundle marginal costs, $\hat{mc}_{jndm}$, and assuming this error is uncorrelated with the size and vertical integration status of firm f.[29]

The Nash assumption also implies the necessary conditions of profit maximizing bundle choice for each firm given the price and bundle choices of its rivals. Our estimation uses a subset of these necessary conditions as moment inequalities. The logic is the same as for 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. We punish candidate parameter estimates if they imply that altering observed bundles are profitable deviations for distributors. Firms may have unobservable information about these decisions which, if left unaddressed, would bias our estimates. We assume that the firm's unobservable information is fixed for a given channel across markets, and sum the profit of changing from observed choices across opposite decisions for a given firm and channel pair. For example, we may see Comcast carry Comedy Central in one market and not in another. Our moment inequality conditions are that the sum of the difference between the observed and deviation profits should be weakly positive.

Because adding or dropping channels is a discrete choice, the implied restrictions are inequalities. We follow the set-up in Pakes et al. (2007). From the Nash assumption, the profits to firm f in market n are higher for its chosen and observed bundles and prices than for alternate bundles:

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

[28] Kagan does not disclose from where it obtains the data measured by τ_c. As these costs are widely considered proprietary business information, it is likely that they are only able to measure them with error.

[29] Shocks to marginal costs include the same unobserved labor costs and advertising rates motivating our choice of instruments. These are likely to depend on idiosyncratic features of market n and are unlikely to be correlated with firms' expansion and integration decisions.

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

$$\Pi_{fndm}((\mathbf{b_{fndm}}, \mathbf{b_{-fndm}}), (\mathbf{p_{fndm}}, \mathbf{p_{-fndm}})) \approx r_{fndm}((\mathbf{b_{fndm}}, \mathbf{b_{-fndm}}), (\mathbf{p_{fndm}}, \mathbf{p_{-fndm}})) + \nu_{fndmb,1} + \nu_{fndmb,2}$$

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

Following Pakes et al. (2007), we define

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

and

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

We assume that for two markets ndm and ndm' and the same firm, $\nu_{fndm,b,b',2} = \nu_{fndm',b,b',2} = \nu_{f,b,b',2}$.

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:

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

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

The optimal pricing condition identifies the cost parameters on its own. Furthermore, in its absence the cost parameters are partially identified. Stacking the three sets of moment conditions together yields our full set of input costs moment conditions[30]:

$$\begin{array}{l} \text{Agg. Input Costs} \\ \text{Nash Pricing} \\ \text{Nash Pricing} \\ \text{Nash Bundling} \end{array} \begin{bmatrix} E_f[\hat{\tau}_{fc}(\eta, \varphi)] - \tau_c \\ \frac{1}{J}\sum_j SZ_{jndm}(\hat{mc}_{jndm} - \sum_{c \in C_{jndm}} \hat{\tau}_{fc}(\eta, \varphi)) \\ \frac{1}{J}\sum_j VI_{jndm}(\hat{mc}_{jndm} - \sum_{c \in C_{jndm}} \hat{\tau}_{fc}(\eta, \varphi)) \\ min(0, \frac{1}{J}\sum_j \Delta r_{fndm}(b_{jndm}, b'; \eta, \varphi) + \Delta r_{fndm'}(b', b_{jndm}; \eta, \varphi)) \end{bmatrix} = 0$$

[30]There are additional moments for the Nash Pricing conditions that we use, but suppress for presentation. These are the covariances between year and tier dummy variables with the difference between implied and predicted marginal cost.

We estimate η and φ by minimizing the empirical analog of these moment conditions, with each weighted equally in the estimation.

5.3 Channel-Distributor Bargaining Parameter Estimation

The unobserved parameters of the bargaining game are each conglomerate and distributor's pair-wise bargaining parameters ζ_{fK}. We use no additional data in identifying the bargaining parameters. They are functions of the estimated cost and demand parameters and the protocol of the bargaining game.

In practice, we choose the values of ζ_{fK} to 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 conglomerate will most often prefer higher input costs, the distributor will always prefer lower input costs. The bargaining model, for a fixed vector of $\zeta_{\mathbf{K}}$, resolves this disagreement. Part of the resolution is due to the bargaining protocol and the respective parties' outside options. The rest is due to the bargaining parameters $\zeta_{\mathbf{K}}$. The estimated input costs are an estimate of the actual resolution point. Therefore, the estimated bargaining powers are the $\zeta_{\mathbf{K}}$ which imply equilibrium input costs from the bargaining model as close as possible to estimated input costs.

Identification of ζ_{fK} relies on two key ingredients. First, we are able to estimate pair-specific input costs. Second, the marginal cost of upstream production is commonly known to be zero. When costs are not observed nor separately estimated, they are not separately identified from the bargaining parameters. The analyst would not know if an input cost is high because marginal cost is high or because the upstream firm's bargaining parameter is high. In this application, because of these two ingredients, we are able to separately identify the bargaining parameters from cost parameters.

The ultimate payoffs for each of the parties involved in bargaining is determined after downstream competition has taken place. When solving for equilibrium input costs, we re-compute, for each potential input cost, the viewership, subscription, and pricing decisions at each stage of the model. These equilibrium quantities determine how much advertising revenue is sold and how much revenue the conglomerate receives from each distributor. We model the advertising revenue as a linear function of household hours watched. We estimate a channel-specific advertising price using Kagan advertising revenue data and Nielsen ratings data. Each channel's estimated advertising price is simply its advertising revenue divided by its average national household rating.

Computing equilibrium input costs is computationally demanding. For both the estimation of the bargaining parameters and the counterfactual, we simplify the computational burden by assuming there is one large market and one small market. We further assume there is one cable distributor for the large market and a separate cable distributor for the small market. There are two "national" satellite providers that compete with the cable operators in each market, but must set the same prices and packages 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. The downstream local market structure is the same as in the estimation, and in reality during the time period of the sample: one cable and two satellite options per market. 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 thousands of markets. The simplification allows a connection to the estimated cost parameters by having different sized distributors while economizing on computational time.

6 Estimation Results

Demand Estimates Table 3 presents estimates of the price sensitivity parameter (α), the impact of income on price sensitivity (π_p), and differences across demographics in tastes for the outside good. The estimated price sensitivity parameter, $\hat{\alpha}$, is -0.29 (0.00) for OLS and -0.50 (0.03) for IV using prices of other firms in the same DMA as the key price instrument.[31] This suggests that our instrumental variables strategy is working as theory would predict.[32]

In markets that offer Basic, Expanded Basic, and Digital Basic cable services, this yields an average own price elasticity for Basic of -4.12, for Expanded Basic of -6.34, for Digital Basic of -13.11, and for Satellite of -5.35.[33] These are on par with most previous estimates in the literature[34] and imply median (mean) margins across the services in our data of 44% (46%), in the range of the estimates of average margins above programming costs of 56% estimated by the FCC (FCC (2009, Table 5)).[35]

Table 4 reports, for each channel in our analysis, information about the distributions of WTP implied by our

[31]We explored using other price instruments, including the prices of the same firm in other markets (used in Crawford (2008)), the total number of subscribers to the firm to which that system belonged (a cost shifter analogous to that used in Section 5.2), and channel dummies (approximating changes in marginal costs due to the inclusion of additional channels). The first yielded a qualitatively similar estimate of price sensitivity ($\hat{\alpha}$ = -0.34 (0.03)), the second a lower but imprecise estimate ($\hat{\alpha}$ = -0.16 (0.31)), and the third a much larger estimate ($\hat{\alpha}$ = -1.09 (0.01)). As all but the last of these models are just-identified, there are no over-identifying restrictions to facilitate testing their validity as instruments. The hypothesis that the channel dummies are orthogonal to the demand error is soundly rejected by a Hansen J-test (p-value = 0.000). The combination of strong theoretical justification and better fit with average industry margins (described below) led us to prefer prices of the same firm in other markets as our price instrument.

[32]We also allowed for the possibility of correlation between the instrument and error by calculating the bounds estimator of Nevo and Rosen (2008). Given the plausible correlations between data and error in our setting and the conditions on the correlations in the data outlined in Nevo and Rosen (2008, p.12), these were only able to say that the true estimate is at least as negative as our preferred IV estimate.

[33]Table 13 in Appendix B.2 reports the full table of own- and cross-price elasticities.

[34]The FCC (2002) (-2.19), the GAO (2003) (-3.22), Beard, Ford, Hill and Saba (2005) (-2.5), Chipty (2001) (-5.9), and Goolsbee and Petrin (2004) (-1.5 for EB, -3.2 for DB, -2.4 for Satellite), have all separately estimated the average own price elasticity of cable services, using market share regressions, diverse data sets, and instrumental variables techniques.

[35]This is a meaningful comparison as we do not impose the restrictions implied by optimal pricing in the demand estimation. Margins are defined as $(p-c)/p$. FCC (2009) estimates total programming expenditure at \$15.8 billion and total Basic, Expanded Basic, and Digital Tier revenue at \$35.6 billion in 2005. 1 - 15.8/35.6 = 56%.

estimates. The first three columns of the table report, for a simulated set of 20,000 households, the mean and standard deviation in WTP for the channel among those that value it positively and the share of households that value it positively. Figure 4 presents estimates of the full marginal distribution of WTP for a subset of these channels.

The WTP estimates mimic the patterns in the Nielsen ratings and Mediamark consumer survey data. The mean and standard deviation of WTP for ESPN ($3.08, $4.46) are higher than for Bravo ($0.65, $0.67) because the mean and variance of ESPN's ratings are higher than Bravo's. The estimated share of households with positive tastes for TNT (0.72) is higher than for the Golf Channel (0.12) because more consumers report watching TNT than the Golf Channel.

The dispersion in WTP for any given channel can be decomposed into the dispersion which can be attributed to demographics and that which cannot. Dispersion due to demographics comes through the impact of demographics on tastes (i.e., Π or $\rho_{d_i c}$) while further dispersion comes through the distribution of unobserved tastes for channels, G. On average across channels, 5% of the dispersion in WTP can be attributed to demographics, although this can be much higher for individual channels.[36] Columns three and four provide an example of demographic effects by reporting mean WTP for family and black households, respectively. Family households are estimated to prefer channels offering family-oriented programming like the Disney Channel and Nickelodeon. Black households are estimated to generally value channels more highly, with a strong effect for BET ($4.54 versus $1.27 among all households).

Correlations in WTP between pairs of channels can arise through demographic groups sharing tastes for those channels, or through the correlations estimated in G. Most pairwise correlations are between -0.1 and 0.1, although some pairs of channels have stronger correlations. We estimate that ESPN and ESPN2 have a correlation in household WTP of 0.67, ESPN and Fox Sports of 0.39, MTV and SoapNet of -0.13, and CNBC and Comedy Central of -0.19. The last column in Table 4 shows that the channel estimated to have the highest correlation in tastes for each channel accords with intuition in who is likely to be the target audience of the programming on both channels.

Input Cost Estimates We estimate median marginal costs for bundles to vary from $11.08 for Basic to $20.74 for Digital Basic packages.

The demand estimates are combined with Nash pricing and bundling assumptions and EBCN average input costs per channel to estimate differences in per-channel input costs across distributors. We 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 with any statistical power. By bringing the extra information contained in EBCN's average costs and the Nash in bundling assumptions, we are able to estimate not only channel specific input costs, but also how those input costs differ for downstream firms based on size and vertical integration.

[36] We calculate this by regressing, for each channel, WTP for the channel among 20,000 simulated households on their demographics and then constructing a weighted average of the R^2 from those regressions using the mean WTP for the channel as a weight.

The estimated input cost parameters, η and φ, in Table 5 imply that Comcast, a distributor with roughly 24 million subscribers, faces input costs 17% below those of a small distributor.[37] The estimated effect of vertical integration is negative and statistically different from zero. Of the three moment conditions, the EBCN average costs help pin down the overall level of input costs while the Nash in pricing and bundling assumptions help pin down how those input costs vary across distributors of different size and/or integration status. For robustness, the second set of columns of Table 5 report the same estimates excluding the Nash in bundle moments conditions. There are few differences.

Most of the patterns in the data generating these estimates are clear from Table 6. It shows that observed prices and estimated marginal costs are lower on average for large distributors, conditional on the characteristics of the bundle. Consequently, we estimate large distributors to have lower per-channel input costs. Prices for bundles are lower for distributors who offer many of their own vertically integrated channels, although we find that estimated marginal costs are not.[38] One might expect these distributors to at least carry their vertically integrated channels more often than other distributors, but this is not true for most of the vertically integrated channels we examine.[39]

Bargaining Parameter Estimates We report our estimates of channel conglomerates' bargaining parameters relative to distributors in Table 7. Smaller values indicate relatively more bargaining power for channels. We estimate that bargaining parameters are usually between 0.25 and 0.75. These estimates discourage as-

[37]We report standard errors using the conservative estimates in Pakes et al. (2007, Section 3.1.3) (PPHI). Andrews and Soares (2010) introduce an alternative procedure to that in PPHI for calculating confidence sets and test statistics that are not asymptotically conservative (and, more generally, have the correct asymptotic size). As our primary results do not depend on hypothesis tests of these parameters and the Andrews and Soares method is more costly to implement, we use the simpler PPHI formulas.

[38]The vertical integration results in both our structural and reduced-form models were sensitive to how we treated outlier values of marginal costs. Sample statistics for the marginal cost estimates for each of our 25,000 bundles had a mean of 9.0 and a standard deviation of 70.7. The standard deviation was so large due to some very small and some very large values (themselves driven by very small and very large market shares). In the analysis, we chose to truncate our estimated costs from below at zero and from above at the price of the bundle. The mean and standard deviation of our truncated costs was 12.0 and 9.1. We found no evidence of effects of vertical integration in the structural analysis with the untruncated costs; the evidence for vertical integration effects reported above is for the truncated costs. The positive and significant vertical integration result in the reduced-form regressions is surprising and due, we suspect, to the difficulty projecting marginal costs onto channel dummies without the restriction that the weighted average across distributors be on par with industry averages reported by EBCN (as in the structural analysis). A median regression of marginal costs on firm size and integration status yields a negative (but statistically insignificant) effect of vertical integration.

[39]Table 14 in Appendix B.2 demonstrates this for the carriage of channels owned by Time Warner between 2004 and 2007. It is true, however, that integrated distributors are more likely to carry their own networks for some new channels that are too small to be included in either the TMS or Nielsen viewing data and are therefore not part of the analysis in this paper. 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. 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 of vertically integrated channels using data from 1991 and finds that integration does affect costs and carriage. Here, we show that this is indeed true if one looks at certain less-established channels, but not for the established channels between 1997 and 2007.

suming take-it-or-leave-it offers as the estimated bargaining parameters are neither zero, which would imply channels take all the marginal surplus, nor one, which would imply distributors do. We estimate that distributors generally have higher bargaining parameters than channel conglomerates for small channel conglomerates (Comcast, Scripps, Rainbow Media, Discovery, Hallmark, Lifetime, Oxygen, Weather Channel, and TV Guide), but that the situation is reversed for large channel conglomerates (ABC Disney, Viacom, NBC Universal, News Corporation, and Time Warner). Among distributors, small cable operators and satellite providers have slightly less estimated bargaining power than large cable operators.

7 The Welfare Effects of À La Carte

7.1 Theoretical Predictions

For a fixed set of channels and ignoring capacity constraints, the socially optimal allocation would deliver every channel in existence to each household that has a positive willingness to pay for that channel. Bundling excludes households that have positive willingness to pay for some channels, but not enough for the full bundle to justify its price. À la carte pricing of channels allows for those excluded under bundling to purchase some channels. However, à la carte partially excludes households who have positive valuations for channels that do not exceed the prices at which the channels are being sold. Which of these two effects dominates determines the total welfare effect of à la carte, and is one output of the counterfactual exercise.

How the surplus generated by multichannel television service is split between and within consumers and firms is also of importance to policy makers. Bundling theory under monopoly suggests that consumers with highly variant preferences, as we estimate television households to be, are better off under ̀la carte pricing in the short run (Adams and Yellen (1976)). The theory under oligopoly is less established and offers ambiguous predictions about the effects of à la carte on consumer welfare. Furthermore, neither of these literatures consider the welfare effects allowing for renegotiation of linear contracts between upstream and downstream firms.

In the long run, the conclusions of economic theory on the welfare effects of à la carte are even less clear. Many opponents of à la carte claim smaller channels appealing to niche tastes will become unprofitable and exit in an à la carte environment. Others claim they may invest less in program quality. We do not model the impact of à la carte on these long-run outcomes. Further research of their evolution in an equilibrium setting is necessary to assess these effects of à la carte regulations.

7.2 Counterfactual Simulations

Supporters have suggested various implementations of à la carte policies. These range from requiring firms which bundle to allow consumers to opt out of programming and receive a rebate (as in the Family and Consumer Choice Act of 2007) to separately priced theme tiers to offering separately priced individual channels. We simulate three outcomes: full à la carte (ALC), theme tiers (TT), and bundle-sized pricing (BSP).

In all our simulations, we make a number of assumptions consistent with a short-run analysis. We assume that preferences are invariant to the policy change. As discussed above, we assume that channels do not alter their programming following the policy change, nor do new channels enter or existing channels exit. We assume the technical, administration, billing, and marketing costs of firms are the same when firms are allowed to bundle as when firms are forced to sell channels à la carte. Finally, we assume that households don't incur any extra cognitive costs from choosing from the larger choice set.

In what follows, we describe in some detail our preferred results. They represent our best estimates of what outcomes would be under various counterfactual policy environments. We recognize, however, that there are many assumptions underlying the specific numbers we present below. In Appendix B.2, we assess the robustness of our conclusions to some of the assumptions underlying our analysis.

Full ALC Our baseline simulation has one large and one small cable market as in the bargaining power estimation. Each is served by its own cable provider and two "national" satellite providers. The demographic distribution for each market is that of the whole United States.

Table 8 summarizes our baseline results. We report economic outcomes implied by our estimates under three scenarios. The first scenario is a bundling equilibrium where each distributor competes by setting a single fixed fee for a bundle of all the 49 channels in our analysis. Table 9 lists the included channels. The second scenario is a Full ALC equilibrium without renegotiation. In this counterfactual, each distributor competes by setting a fixed fee and separate à la carte prices for each channel in the specification. The input costs they face do not allow for renegotiation, however. That is, the input costs are the same as those we estimate in a world with only bundles. While unrealistic in television markets, this is the maintained assumption in most of the theory literature analyzing this issue. The last scenario is again Full ALC, but allows for the renegotiation of input costs taking as given the bargaining parameters we estimate for each channel conglomerate-distributor pair.[40]

We also simulate the effects of ALC on channels' advertising revenue. For each channel, we assume that the price per minute of advertising they receive under bundling will also be what they receive under ALC.

[40] In this equilibrium, we made the simplifying assumption that distributors set ALC prices equal to their agreed-upon input costs and earned profits only on fixed fees for access to their platforms. We did so for computational reasons. Solving for renegotiated input costs in the full ALC equilibrium requires repeatedly solving for downstream prices at candidate input costs. Numerical errors in those pricing equilibria appear to propagate into the bargaining equilibria at tractable convergence tolerances, making that optimization non-smooth. It also makes it extremely time-consuming as the pricing equilibria must be repeated at each iteration in the solution of the input costs for each distributor-conglomerate pair and these in turn must be iterated to obtain the bargaining equilibrium. We feel comfortable with this assumption for two reasons. First, before imposing it we were finding downstream markups of between -5 and 10% for input costs close to but not quite reaching equilibrium values. Second, it is consistent with the predictions of Armstrong and Vickers (2001) and Rochet and Stole (2002) who find cost-based two-part tariffs characterize the equilibria in some settings analyzing competition among price-discriminating firms. In Appendix B.2, we allow for downstream margins to be 10% rather than 0 and obtain qualitatively similar results.

The change in their advertising revenue is then simply given by their current advertising revenue times the percentage change in their viewing implied by the counterfactual. This is converted to a per-household basis when calculating total revenue in Tables 8 and 9.

The top panels in Table 8 present general features of the various equilibria. We see that while most households purchase some cable or satellite service in the bundling equilibrium, this is even greater under à la carte as households unwilling to pay the full cost of the bundle opt to purchase a smaller number of channels. As expected, households under ALC purchase fewer than the full complement of channels.

The bottom panels in Table 8 summarize the welfare effects of ALC. Comparing first the bundling and Full ALC *without* renegotiation, we see that channel profits drop significantly (despite an increase in advertising revenue), distributor profits increase slightly, and overall industry profits fall (by 12.7%). Consistent with the theory literature, consumer surplus rises by 19.2%, driven both by reduced expenditure among those that previously purchased the bundle and the addition of households that were previously excluded from the market. The increase in consumer surplus outweighs the fall in profits, meaning total surplus rises by 4.1%.

Allowing for renegotiation in the last set of columns changes these conclusions. Most input costs increase, some dramatically so. The total for the channels in our analysis increases by an estimated distributor-share-weighted average of 103.0%, increasing prices paid by households. Mean consumer expenditure increases an estimated 2.2%.

These input cost increases also have important effects on welfare. Instead of reducing channel profits, all of channel, distributor and industry profits are estimated to increase, the latter by 4.8%.[41] Estimated consumer surplus is effectively unchanged (+0.2%). The predicted change in total welfare is still positive, but lower than before renegotiation as some households no longer purchase some channels of moderate value whose input costs and thus prices rise.

Table 9 breaks down the input cost and profit effects by the channels included in our analysis.[42] The first three columns report the estimated share-weighted monthly license fee per subscriber under bundling, the license fee under ALC with renegotiation, and the percentage change. There is considerable heterogeneity across channels in the effects of ALC. Some channels are estimated to increase their license fees by 300% or more (Animal Planet, Food Network, TV Land), while others are estimated to cut their fees (Nickelodeon, Oxygen, TV Guide).

There are similarly heterogeneous effects on channel revenues. The remaining columns in Table 9 report

[41] This need not be surprising. There is tremendous uncertainty in the industry about outcomes in an ALC world. Neither channel nor distributors may know the structure of demand for channels and/or bargaining outcomes under ALC. Our results suggest ALC would be profitable for the industry. Of course, any equipment, administration, billing, or marketing costs arising under ALC would reduce these profits, further reducing consumer surplus and likely causing total surplus to fall.

[42] The results described in this table should be interpreted under the maintained assumption that the more households watch a channel, the more they value that channel. In Appendix B.2, we conduct a monte carlo analysis to explore the consequences of allowing channels that are watched less by households to nonetheless be valued more (and vice versa) and find that it may yield underestimates of WTP for channels for which household tastes are high for early minutes but decline quickly with minutes watched (e.g. sports programming) and overestimates of WTP for channels for which household tastes are more constant across minutes. See the Robustness subsection below and in Appendix B.2 for more detail about this issue.

total (license fee plus advertising) per-household revenue to each channel under bundling and ALC with renegotiation, the change between them, and the percentage change in the component (license fee, advertising) revenues. Total channel affiliate fee revenue decreases by an estimated 3.7% and advertising revenue increases by 10.1%, the latter driven by increased viewership by households that did not purchase under bundling. There is significant estimated heterogeneity across channels, with some predicted to lose 40% or more of their revenue (GSN, Oxygen, Versus) while others are predicted to increase revenue by 100% or more (Animal Planet, CNN, History Channel).

Theme Tiers and Bundle Sized Pricing We also simulated two alternative regulatory scenarios. In the Bundle-Sized Pricing (BSP) scenario (Chu et al. (2010)), we assume downstream firms continue to offer a bundle of all the channels, but add to this a package of fifteen channels assembled by each household according to their tastes. In the Theme-Tier (TT) scenario, we assume downstream firms offer five tiers of service (Sports, News, Family and Education, Music and Lifestyle, and General) from which a household can choose any combination.[43] In this scenario, distributors also charge a fixed fee. In both scenarios, distributors and channel conglomerates renegotiate input costs. Table 10 reports the results.

Outcomes under both BSP are TT are worse for consumers. In each case, input costs are estimated to rise almost as much as under Full ALC, but consumer choice is more restricted, reducing their benefits. Under BSP, consumers are able to choose their 15 favorite channels (and many do), but pay a similar amount to Full ALC while getting fewer channels. This reduces their consumer surplus (by 8.8%). Total industry profit is similar and total surplus falls (by 2.3%). Outcomes under theme tiers are more dramatic. Households watch as many channels as Full ALC, but now pay much more to do so (consumer expenditure increases an estimated 33.8%). Estimated consumer surplus therefore falls considerably (-22.0%). Channel profits soar, yielding an aggregate predicted industry profit increase of 24.2%. Total surplus is effectively unchanged (-0.2%) relative to the bundling baseline.

Results Summary Our findings confirm the intuition regarding the likely effects of ALC described in Section 2. When we do not allow for renegotiation (Table 8, Columns 2-3), we turn off the input-cost-raising bargaining effect and find consumer surplus increases considerably (+19.2%) and industry profits fall (-12.7%). As suggested by much of the bundling literature, for fixed input costs, we find bundling transfers surplus from consumers to firms. When we allow for renegotiation (Table 8, Columns 4-5), costs rise (+103.0%), prices follow suit, and these consumer surplus gains are effectively eliminated (+0.2%). Things are even worse for consumers under bundle-sized pricing and theme tiers (Table 10, Columns 4 & 7). The bundling of channels within each of these alternatives eliminates much of the consumer surplus benefits accruing under Full ALC and *still* almost doubles input costs. This worst-of-both-worlds outcome significantly lowers consumer surplus (by 8.8% or 22.0%). Our qualitative conclusion is that consumers

[43] See the notes to Table 10 to see the identities of the channels included in each tier.

could in principle benefit from mandatory à la carte at existing input costs, but would not in practice benefit due to input cost renegotiation in an à la carte world.

Robustness Our goal is to accurately measure the welfare effects of à la carte pricing in multichannel television markets. As such, it is important to have confidence that this fundamental conclusion is robust and not sensitive to particular assumptions underlying the model, estimation, or counterfactual simulations. In Appendix B.2, we consider the robustness of our results to alternative assumptions on demand, cost, and bargaining, including allowing for positive channel margins for distributors in the counterfactual, different distributional assumptions for preferences, turning off unobserved correlation in tastes, and allowing renegotiated input costs to be half or double what we estimate. We also conduct a monte carlo exercise in a simplified economic environment to explore the likely consequences of relaxing our assumption that a channel which is watched more is necessarily valued more.

Table 16 in Appendix B.2 shows that alternative assumptions about the downstream margins and the shape of and correlation between household preferences for channels yield qualitatively similar results: estimated consumer surplus changes between -5.4% and 0.2%, profits between 2.4% and 12.8%, and total surplus between -1.7% and 6.0%. Bargaining outcomes are much more important for predicting surplus: if renegotiated input costs were to rise by half (double) the 103.0% we estimate, estimated consumer surplus would increase by 18.5% (fall by 27.6%). This merely emphasizes the importance of estimating a bargaining game and simulating counterfactual bargaining outcomes in order to accurately understand the effects of unbundling in television markets.

Relaxing the assumption that channels that are watched more are valued more in our monte carlo exercise yielded interesting insights. Table 17 in Appendix B.2 shows that a range of channel-specific economic outcomes are mis-estimated when households watch some channels less but nonetheless value them more. In particular, WTP, prices, and market shares for these channels are underestimated while the same outcomes for those that are watched more but valued less are overestimated. Adding across channels, however, causes these errors to cancel out and, in the monte carlo, yields statistically similar predictions for the overall welfare effects of à la carte policies.

8 Conclusion

This paper has combined a structural model of the multichannel television industry with market and viewership data in order to evaluate the welfare effects of proposed à la carte pricing regulations. We extend a standard demand model to a setting of joint purchasing and viewership decisions and combine it with a model of distributor pricing and bundling, and channel-distributor bargaining. We estimate the model using demand, pricing, viewership, and cost data from the industry. We use the estimated model to simulate an unrealized regulatory environment: à la carte pricing regulations. Critically, we allow for the renegotiation of supply contracts under à la carte and find that total input costs for the 49 channels in our analysis would rise by 103.0%. We compare the distributions of consumer and producer surplus under a simulated bundling setting with those under à la carte allowing for these cost increases and predict that, in the short run, consumer welfare would change between -5.4% and 0.2% under à la carte regulations, while industry profits and total surplus would increase between 2.4% and 12.8% and -1.7% and 6.0%, respectively. Any implementation or marketing costs of à la carte could make it worse for all.

One could improve our analysis of bundling in the multichannel television industry in future work by trying to relax some of the most important maintained assumptions in our analysis. Relaxing the assumption that households value equally time spent watching different channels, allowing for asymmetric information in channel-distributor bargaining, and analyzing for the long-run effects of à la carte regulations on entry, exit, and the content and quality of channels would all be valuable.

A The Multi-Channel Television Industry

The multi-channel television market is a two-sided market. Cable and satellite systems provide a platform connecting households with both program producers and advertisers. Figure 5 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.[44] 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, we focus on the for-pay distribution and advertising markets.

Cable television systems choose a portfolio of television channels, bundle them into services, and offer these services to consumers in local, geographically separate, markets. Satellite television systems similarly choose and bundle channels into services, but offer them to consumers on a national basis.

All cable and satellite systems offer four main types of channels. *Broadcast channels* are advertising-supported television signals broadcast over the air in the local cable market by television stations and then collected and retransmitted by cable systems. Examples include the major, national broadcast channels – 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 distributed nationally to systems via satellite. Examples include MTV, CNN, and ESPN. *Premium programming channels* are advertising-free entertainment channels. Examples include HBO and Showtime. *Pay-Per-View* are specialty channels devoted to on-demand viewing of the most recent theatrical releases and specialty sporting events.

Broadcast channels and cable channels are typically bundled and offered as *Basic Service* while premium programming channels are typically unbundled and sold as *Premium Services*.[45] Distributors now offer cable channels on multiple services, called *Expanded Basic* and *Digital Services*.

Most advertising space is sold by channels, but also for a few minutes per hour by the local cable system.[46] Advertising revenues account for nearly one half of total channel revenues. Advertising revenues depend on the total number and demographics of viewers. These figures, called ratings, are measured by Nielsen Media Research (hereafter Nielsen). Ratings are measured at the Designated Metropolitan Area (DMA) level, of which there are 210 in the United States. In urban areas, the DMA corresponds to the greater metropolitan area. DMA's usually include multiple cable systems with different owners.

[44] 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 placed in its programming lineup. These channels are then distributed to consumers in one of two 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 systems 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 households via the Internet.

[45] 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.).

[46] Local advertising revenue to cable systems for 2006 accounted for approximately 5% of total cable system revenue.

Figure 1: Dispersion in WTP for components is higher than dispersion in WTP for a bundle

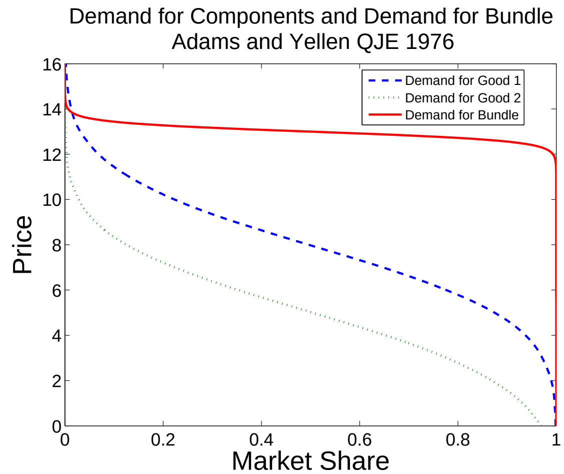
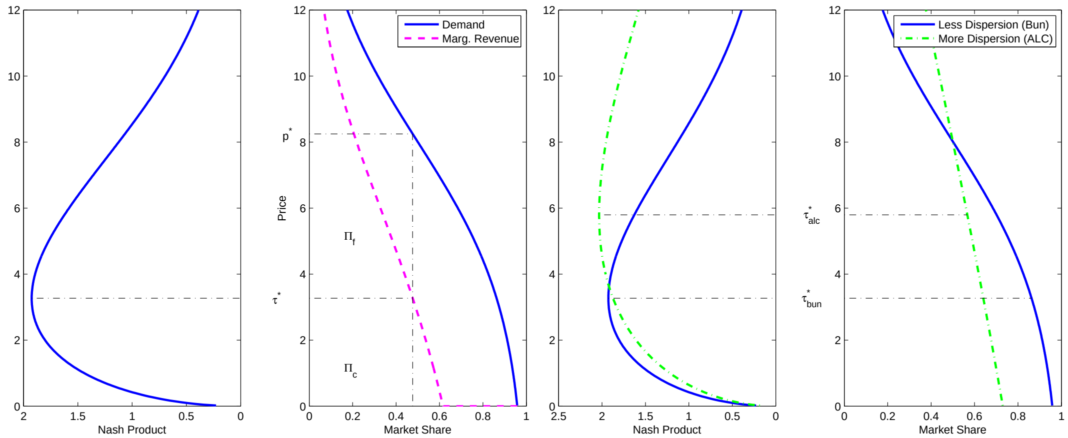


Figure 2: Nash Bargaining for Input Costs



Notes: These figures provide the intuition for the determination of input costs under Nash Bargaining. The left figure shows the value for the input cost that maximizes the Nash Product under bilateral monopoly with linear fee contracts and symmetric bargaining parameters. The solid lines in the right panel of the left figure show the demand and marginal revenue for the product faced by the downstream firm. Total (gross) profit is divided between the downstream distributor (π_f) and the upstream content providers (π_c) according to an input cost (τ). The marginal cost to the content provider is assumed to be zero. The left panel of the left figure reports the value of Nash Product (as in Equation (7) for different values of τ. The reported input cost maximizes the Nash Product.

The right figure demonstrates the consequences to input costs of the firm facing a product with more dispersion in tastes (as typically happens under à la carte pricing). At the optimal input price in the left figure, the downstream firm wishes to raise price and earns a greater share of the total profit. The upstream content provider recognizes this and bargains for a higher input cost. These dynamics are evident in the shape of the Nash Product for the more dispersed tastes.

Table 1: Factbook Summary Statistics

	N	Mean	SD	Min	Max
All Bundles					
Price	25,490	23.46	9.20	0.00	87.06
Market Share	25,490	0.44	0.27	0.00	0.99
Total Cable Channels	25,490	20.3	16.1	0	176
Basic Only Markets					
Basic Service					
Price	14,732	23.70	6.36	0.00	80.25
Share	14,732	0.54	0.22	0.00	0.99
Total Cable Channels	14,732	17.3	9.4	0	95
Basic and Exp. Basic Markets					
Basic Service					
Price	4,046	13.49	5.71	0.00	47.67
Share	4,046	0.11	0.15	0.00	0.89
Total Cable Channels	4,046	8.91	7.68	0	56
Expanded Basic Service					
Price	4,046	27.39	7.92	0.00	87.06
Share	4,046	0.57	0.19	0.00	0.97
Total Cable Channels	4,046	26.5	10.0	0	77
Basic, Exp. Basic, and Dig. Basic Markets					
Basic Service					
Price	493	13.26	5.60	0.00	38.68
Share	493	0.09	0.09	0.00	0.65
Total Cable Channels	493	8.3	6.3	1	35
Expanded Basic Service					
Price	493	34.62	7.81	0.00	61.51
Share	493	0.39	0.16	0.01	0.84
Total Cable Channels	493	47.1	10.7	18	89
Digital Basic Service					
Price	493	44.56	10.07	0.00	70.27
Share	493	0.15	0.10	0.00	0.53
Total Cable Channels	493	78.8	19.1	37	176

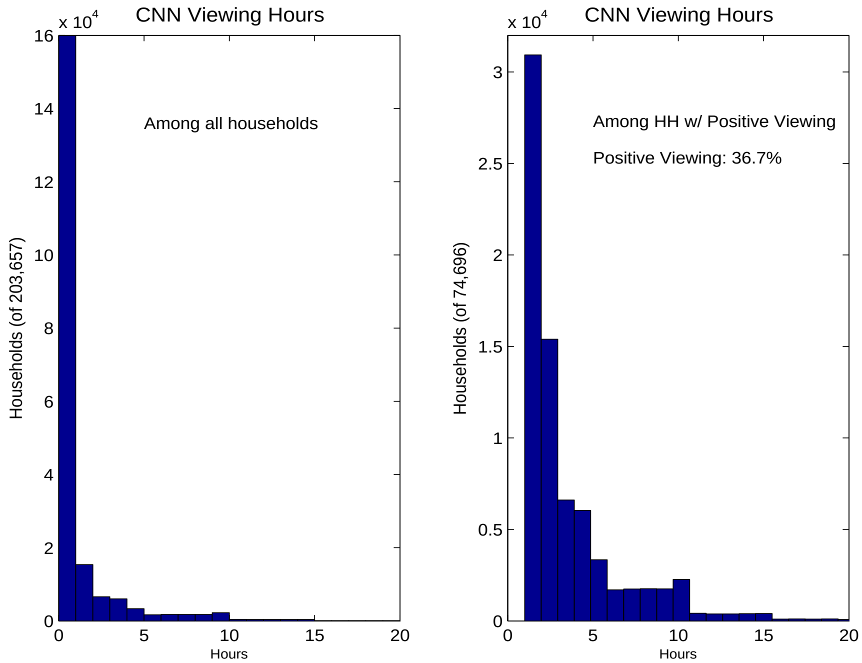
Notes: This table reports sample statistics from our individual cable system (Factbook) data for all markets and by type of bundles they offer. An observation is a system-bundle-year. Prices are in 2000 dollars. Market shares are defined as subscribers divided by homes passed, with homes passed defined as the set of households able to purchase cable service from each system. Both are in the data. Total cable channels is the sum of over 350 television channels carried by cable systems in the Factbook.

Table 2: Channel Summary Statistics

	Cable System Carriage		Household Viewership			
Data Source	Factbook		Nielsen	Mediamark		
Channel	Any Tier (Pcntge)	Basic Tier (Pcntge)	Mean Rating	Mean Rating	StdDev Rating	Cume
ABC Family Channel	91.2	75.7	0.4	0.6	1.5	31.6
AMC	55.3	30.9	0.5	0.6	1.4	27.2
Animal Planet	22.8	12.1	0.3	0.6	1.5	34.8
Arts & Entertainment	68.3	48.7	0.7	0.8	1.7	37.8
BET Networks	21.1	10.9	0.4	0.3	1.5	10.6
Bravo	13.3	3.3	0.2	0.2	0.7	14.4
Cartoon Network	29.1	15.7	1.6	0.5	1.8	20.9
CNBC	37.6	19.7	0.2	0.5	1.4	29.5
CNN	94.5	77.5	0.7	1.8	3.0	53.8
Comedy Central	25.1	11.1	0.5	0.5	1.3	27.6
Country Music TV	48.0	37.2	0.2	0.2	1.0	13.5
Court TV	16.2	4.5	0.4	0.4	1.4	18.1
Discovery Channel	88.0	71.6	0.6	1.1	1.9	50.9
Disney Channel	41.6	29.6	1.2	0.5	1.4	21.2
E! Entertainment Television	22.9	11.0	0.3	0.3	0.9	24.4
ESPN	96.7	76.7	0.9	1.1	2.2	40.7
ESPN 2	36.6	21.4	0.3	0.5	1.4	25.2
Food Network	13.6	4.5	0.4	0.5	1.5	26.7
Fox News Channel	20.0	10.0	0.8	1.0	2.2	40.0
Fox Sports Net	19.4	11.3	0.3	0.4	1.2	20.2
FX	21.0	9.9	0.5	0.4	1.2	23.3
GSN	8.7	0.8	0.2	0.2	0.9	7.4
Golf Channel	10.9	1.8	0.0	0.1	0.6	6.9
Hallmark Channel	8.2	3.3	0.3	0.2	1.0	10.8
HGTV	26.3	13.2	0.6	0.6	1.6	27.5
History Channel	32.0	18.5	0.6	0.8	1.7	37.9
Lifetime	63.2	41.8	0.9	1.0	2.2	34.4
MSNBC	14.4	5.0	0.3	0.5	1.3	30.2
MTV	52.7	30.2	0.7	0.4	1.4	21.8
MTV2	1.9	0.1	0.0	0.1	0.7	7.8
National Geographic Channel	6.5	1.1	0.1	0.2	0.8	13.2
Nickelodeon	73.8	52.5	1.8	0.4	1.3	17.7
Oxygen	2.8	0.2	0.1	0.1	0.5	7.2
Syfy	33.4	18.4	0.5	0.4	1.4	20.9
SoapNet	4.0	0.4	0.1	0.1	0.6	2.5
Speed Channel	11.8	3.2	0.1	0.1	0.7	7.8
Spike TV	24.0	15.0	0.5	0.4	1.1	18.9
TBS Superstation	96.3	90.7	1.1	0.9	1.7	39.8
The Weather Channel	64.1	46.0	0.3	0.7	1.3	50.3
TLC	45.1	29.9	0.5	0.5	1.3	29.0
TNT	85.2	63.7	1.3	0.9	1.8	41.3
Toon Disney	8.6	2.1	0.2	0.1	0.7	6.1
Travel Channel	16.8	8.3	0.2	0.2	0.7	18.7
TV Guide Channel	19.3	11.5	0.2	0.2	0.6	17.5
TV Land	23.2	15.0	0.8	0.6	1.8	23.9
USA Network	88.8	66.3	1.2	0.8	1.6	37.4
Versus	9.3	1.4	0.1	0.1	0.5	4.8
VH1	39.6	22.6	0.4	0.3	0.9	18.2
WE: Women's Entertainment	7.2	0.8	0.1	0.1	0.5	5.9

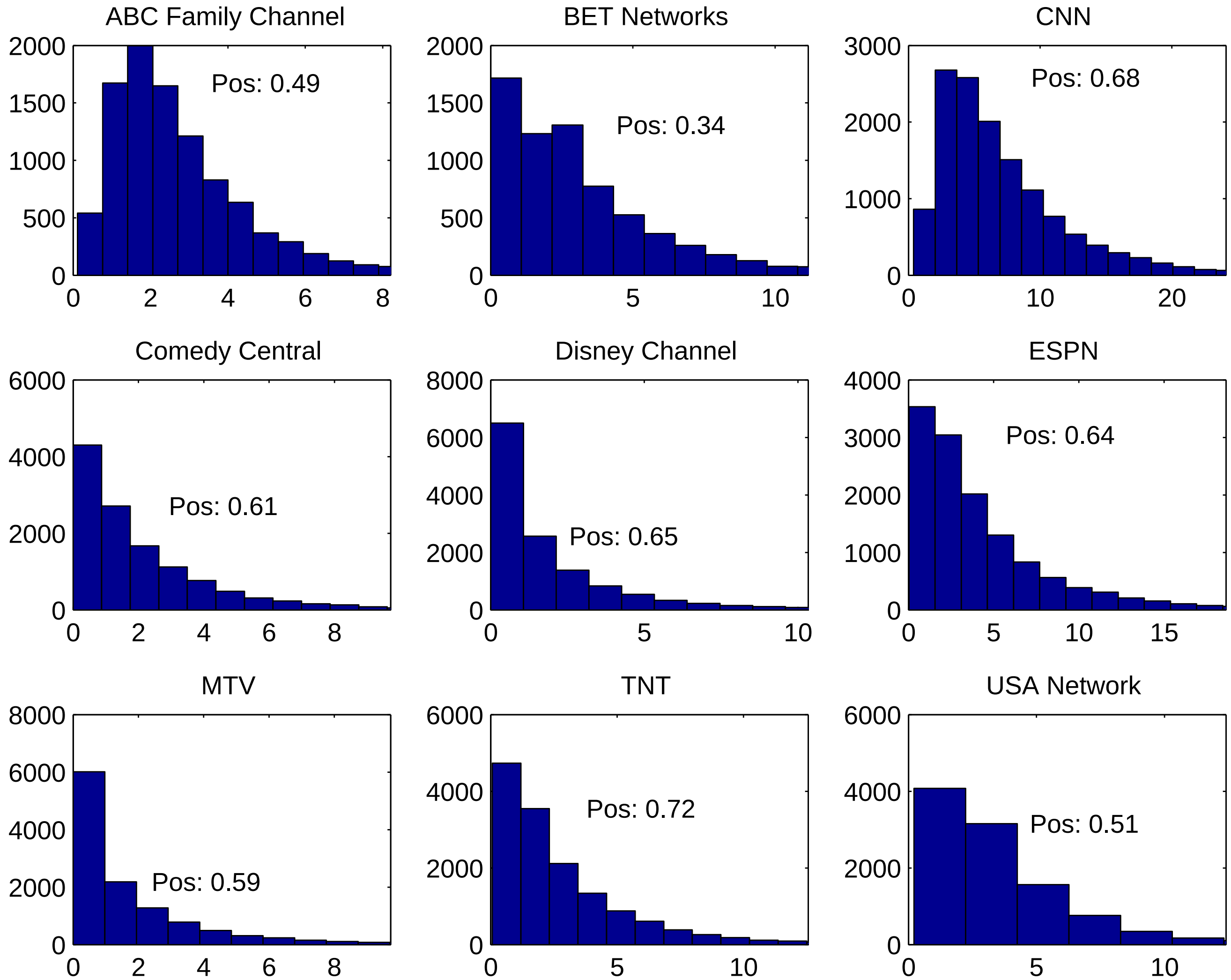
Notes: This table reports summary statistics for channels from both our cable system (Factbook) and viewership (Nielsen, Mediamark) data. The channels reported are those cable channels for which we could get complete data from all three channel data sources used in our analysis. The first column reports the average carriage of each cable channel on any offered tier of service across our system-years. The second column reports average channel carriage on just the Basic tier. The last four columns report summary statistics about household viewing patterns across channels from our Nielsen and Mediamark data. The third column reports the average rating for all programs on that channel for the four Nielsen sweeps months (Feb, May, Aug, Nov) between 2000 and 2006. The fourth and fifth columns report the mean and standard deviation of the fraction of households reporting viewing each channel per hour for our sample of Mediamark households from 2000 to 2007. This is analogous to an average Nielsen rating for that channel and we therefore call them "ratings" above. The last column reports the fraction of Mediamark households reporting positive viewing for each channel. This is known as the channel's "cume," short for cumulative audience.

Figure 3: Distribution of Viewing for CNN, Mediamark (MRI) Data



Notes: This figure reports the distribution of viewing hours reported by our 200,000+ MRI households for CNN. The left panel shows the distribution of viewing for all MRI households, including the 63.3% that report no viewing. The right panel shows the distribution of viewing among the 36.7% of households that report positive amounts of viewing. Note the positive skewness in the distribution; similar patterns arise for all channels. This motivates our assumption that the marginal distributions of unobserved tastes for channels follows a mixture distribution with a mass point at zero and an exponential distribution among those with positive values.

Figure 4: Estimated WTP for a Subset of Channels



Notes: This figure documents the estimated willingness-to-pay for a subset of cable channels among 20,000 simulated households. Reported is the share of those households that value each network positively and the distribution of WTP among that subset. In each figure, the y-axis reports households and the x-axis reports WTP in 2000 dollars.

Table 3: Price Sensitivity and Non-Television Preference Parameters

Parameter	*Estimate*	*Standard Error*
Price Sensitivity (IV)	-0.50	0.03
Price Sensitivity (OLS)	-0.29	0.00
Price Income Interaction	0.11	0.01
Family x Outside Good	0.00	0.04
Income x Outside Good	0.64	0.17
Black x Outside Good	0.70	0.24
Hispanic x Outside Good	3.97	4.11
Asian x Outside Good	3.24	1.92
Bachelors x Outside Good	2.45	0.36
Age x Outside Good	1.07	0.29

Notes: This table reports our GMM results for a subset of demand parameters, including the estimated mean marginal utility of income, α, the impact of income on marginal utility, π_{yp}, and differences across demographics in tastes for the outside good. Also reported is the estimated mean marginal utility from the same estimation procedure without price instruments, which we denote OLS.

Table 4: Estimated WTP

Channel	Mean WTP	StdDev WTP	Share Positive	Mean WTP Family HH	Mean WTP Black HH	Highest Correlated Channel
ABC Family Channel	1.59	2.24	0.49	1.68	1.80	'TV Land '
AMC	1.40	1.59	0.51	1.15	1.83	'MSNBC '
Animal Planet	2.05	3.02	0.58	2.08	1.81	'National Geographic Channel '
Arts & Entertainment	2.10	2.63	0.58	1.90	2.23	'History Channel '
BET Networks	1.27	2.74	0.34	1.34	4.54	'MTV2 '
Bravo	0.65	0.67	0.61	0.63	0.76	'ESPN '
Cartoon Network	2.06	4.01	0.49	2.27	2.54	'Nickelodeon '
CNBC	2.02	2.97	0.55	1.84	2.01	'CNN '
CNN	5.38	5.91	0.68	4.94	8.30	'Fox News Channel '
Comedy Central	1.51	2.39	0.61	1.52	1.34	'MTV '
Country Music TV	0.89	1.56	0.57	0.89	0.79	'Food Network '
Court TV	1.76	3.11	0.50	1.79	2.23	'Arts & Entertainment '
Discovery Channel	2.70	2.99	0.65	2.55	2.67	'Animal Planet '
Disney Channel	1.43	2.51	0.65	1.52	1.72	'Nickelodeon '
E! Entertainment Television	1.15	1.69	0.62	1.16	1.10	'VH1 '
ESPN	3.08	4.46	0.64	2.86	3.63	'ESPN 2 '
ESPN 2	1.80	3.12	0.62	1.75	2.02	'ESPN '
Food Network	2.06	3.25	0.71	2.08	2.18	'TV Guide Channel '
Fox News Channel	4.07	5.89	0.60	4.10	4.69	'CNN '
Fox Sports Net	1.63	2.82	0.55	1.58	1.55	'ESPN 2 '
FX	1.45	2.59	0.51	1.47	1.41	'USA Network '
GSN	0.74	2.97	0.08	0.83	1.51	'ESPN 2 '
Golf Channel	0.52	1.86	0.12	0.38	0.68	'CNN '
Hallmark Channel	1.43	3.96	0.16	1.47	2.09	'Country Music TV '
HGTV	2.60	4.67	0.42	2.59	3.02	'Food Network '
History Channel	2.70	4.06	0.40	2.53	3.09	'Arts & Entertainment '
Lifetime	2.25	3.73	0.31	2.46	5.57	'AMC '
MSNBC	1.69	3.23	0.29	1.38	2.61	'AMC '
MTV	1.22	2.28	0.59	1.25	1.36	'VH1 '
MTV2	0.71	1.23	0.52	0.79	0.63	'VH1 '
National Geographic Channel	1.03	1.60	0.69	1.04	0.92	'Animal Planet '
Nickelodeon	1.31	2.55	0.50	1.45	1.35	'Disney Channel '
Oxygen	0.41	0.44	0.60	0.49	0.64	'Disney Channel '
Syfy	1.74	2.97	0.54	1.74	1.82	'USA Network '
SoapNet	0.49	1.04	0.42	0.52	0.58	'TBS Superstation '
Speed Channel	0.33	0.41	0.56	0.41	0.19	'Versus '
Spike TV	1.18	2.00	0.57	1.18	1.07	'The Weather Channel '
TBS Superstation	2.05	2.85	0.69	1.98	2.23	'TNT '
The Weather Channel	1.71	1.83	0.70	1.59	1.66	'Spike TV '
TLC	1.82	2.81	0.61	1.84	1.57	'Discovery Channel '
TNT	2.36	3.10	0.72	2.31	2.54	'USA Network '
Toon Disney	0.44	1.69	0.13	0.57	0.90	'Cartoon Network '
Travel Channel	0.76	2.27	0.15	0.80	0.74	'Nickelodeon '
TV Guide Channel	0.50	0.75	0.57	0.54	0.60	'Food Network '
TV Land	2.06	3.40	0.59	2.11	2.45	'ABC Family Channel '
USA Network	2.12	3.19	0.51	2.19	2.62	'TNT '
Versus	0.23	0.31	0.49	0.28	0.21	'Speed Channel '
VH1	0.74	1.28	0.56	0.75	0.90	'MTV2 '
WE: Women's Entertainment	0.45	0.69	0.50	0.49	0.53	'National Geographic Channel '

Notes: This table reports information of the distribution of WTP for channels implied by our estimates. The first two columns report the mean and standard deviation in WTP for each channel among those that value it positively. The third column reports the estimate share of households that do so. The fourth and fifth columns report estimated WTP among family and black households. The last column reports the channel estimated to have the highest correlation in WTP for each channel. WTP is measured in year 2000 dollars per month per household.

Table 5: Input Cost Parameters

	All Moments		No Bundling Moments	
Parameter	*Estimate*	*Standard Error*	*Estimate*	*Standard Error*
Constant	0.16	0.00	0.20	0.00
Kagan Scale	0.91	0.00	0.93	0.00
MSO Size	-0.08	0.00	-0.10	0.00
Vertical Integration Dummy	-0.14	0.01	-0.16	0.01

Notes: This table reports the impact of various factors on our estimated input costs. Kagan scale refers to the input cost for that channel as estimated by Kagan World Media (2008). Distributor (MSO) size is measured in tens of millions of households. Vertical integration is the share of the channel owned by that distributor (between 0 and 1).

Table 6: 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	-0.0955	0.0079	-12.12	-0.055	0.0107	-5.10
Number of Integrated Channels	-0.1668	0.0684	-2.44	0.473	0.093	5.07
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	25490			25490		
R-squared	0.563			0.169		
F(271, 25218)	111.92			18.98		

Notes: This table reports the results of regressions designed to highlight the identification of our input cost estimates. The first set of columns reports the results of a regression of bundle prices on the size of the distributor offering the bundle and a sum of the number of vertically integrated channels in the distributor's bundle. We condition on various variables that might affect marginal costs. The second set of columns reports the results of a regression of our estimated bundle marginal costs on the same covariates.

Table 7: Conglomerate Bargaining Parameters

Conglomerate	Big Cable	Small Cable	DirecTV	Dish Network
ABC Disney	0.28	0.25	0.18	0.17
Viacom	0.49	0.48	0.54	0.53
NBC Universal	0.50	0.49	0.52	0.51
Comcast (Content Division)	0.69	0.68	0.67	0.66
Scripps	0.55	0.55	0.58	0.58
News Corporation	0.42	0.39	0.34	0.32
Rainbow Media	0.70	0.69	0.68	0.67
Discovery Networks	0.62	0.61	0.63	0.63
Time Warner	0.40	0.38	0.38	0.37
Hallmark	0.69	0.69	0.71	0.71
Lifetime	0.43	0.43	0.43	0.43
Oxygen	0.73	0.72	0.71	0.70
Weather Channel	0.69	0.69	0.69	0.69
TV Guide	0.77	0.77	0.76	0.76

Notes: This table reports our estimated bargaining parameters for channel conglomerates versus distributors of various types. Smaller values of the bargaining parameters indicate relatively more bargaining power for channels. Channel conglomerates are ABC Disney (ABC Family Channel, Disney Channel, ESPN, ESPN2, Soap Net, Toon Disney), Viacom (BET Networks, Comedy Central, Country Music TV, GSN, MTV, MTV2, Nickelodeon, Spike TV, TV Land, VH1), NBC Universal (Arts & Entertainment, Bravo, CNBC, MSNBC, Syfy, USA Network), Comcast (E! Entertainment Television, Golf Channel, Versus), Scripps (Food Network, HGTV), News Corporation (Fox News Channel, Fox Sports Net, FX, National Geographic Channel, Speed Channel), Rainbow Media (AMC, WE: Women's Entertainment), Discovery Networks (Animal Planet, Discovery Channel, History Channel, TLC, Travel Channel), Time Warner (Cartoon Network, CNN, Court TV, TBS Superstation, TNT). Hallmark, Lifetime, Oxygen, Weather Channel, and TV Guide are single-channel "conglomerates." See the end of Section 5 for descriptions of the distributor types.

Table 8: Baseline Counterfactual Results: Full À La Carte

	Bundling	ALC No Reneg	% Change	ALC With Reneg	% Change
Non-welfare Outcomes					
Cable & Sat Penetration	0.880	0.998	13.3%	0.993	12.8%
Total Affiliate Fees	$18.22	$18.22	0.0%	$36.98	103.0%
Mean Consumer Expn	$27.63	$21.07	-23.8%	$28.24	2.2%
Number Channels Received	42.8	22.0	-48.5%	19.3	-54.9%
Number Channels Watched	22.2	22.0	-0.5%	19.3	-12.8%
Welfare Outcomes					
Channel Profits					
Total License Fee Rev	$16.03	$7.95	-50.4%	$15.44	-3.7%
Total Advertising Rev	$13.38	$14.71	10.0%	$14.73	10.1%
Total Channel Revenue	$29.41	$22.67	-22.9%	$30.16	2.6%
Distributor Profits	$11.59	$13.11	13.1%	$12.81	10.4%
Total Industry Profits	$41.00	$35.78	-12.7%	$42.97	4.8%
Mean Consumers Surplus	$45.82	$54.59	19.2%	$45.91	0.2%
Total Surplus	$86.82	$90.37	4.1%	$88.88	2.4%

Notes: This table reports the results of our baseline counterfactual simulations of full à la carte (ALC) pricing policies on prices and welfare. The economic environment consists of one large and one small cable market (served by one large and one small cable operator) and two "national" satellite providers, each offering access to their platform and approximately 50 cable channels. In the bundling equilibria reported in column one, each firm competes by pricing a single bundle of channels. In both ALC equilibria, each firm competes by setting a fixed fee and then separate prices for each offered channel. Columns two and three report results for ALC *without* allowing input market renegotiation (i.e. with input costs at their values in the bundling equilibrium); columns four and five allow renegotiation. In the renegotiation equilibrium, we impose that downstream prices equal the renegotiated input costs. See footnote 40 in the text for details. Average outcomes (e.g. Total Affilate Fees, Number of Channels) are weighted across distributors according to their estimated market shares. Dollar values are 2000 dollars per U.S. television household per month.

Table 9: Input Cost and Welfare Effects by Channel

	Input Cost Effects			Profit Effects				
Channel	Bundling Input Cost	ALC Input Cost	% Change	Total Bundling Revenue	Total ALC Revenue	% Change	% Change License Fee Rev	% Change Advert Rev
ABC Family Channel	$0.32	$0.83	156.9%	$0.46	$0.58	24.5%	29.9%	15.9%
AMC	$0.32	$0.54	67.8%	$0.41	$0.43	3.9%	-2.2%	16.9%
Animal Planet	$0.20	$0.97	372.8%	$0.25	$0.53	109.3%	150.0%	9.8%
Arts & Entertainment	$0.31	$1.08	250.6%	$0.57	$0.91	58.8%	109.4%	13.3%
BET Networks	$0.26	$0.58	127.3%	$0.56	$0.55	-1.7%	-26.8%	15.4%
Bravo	$0.27	$0.51	92.3%	$0.39	$0.40	1.4%	2.0%	0.6%
Cartoon Network	$0.26	$0.78	199.1%	$0.54	$0.62	14.7%	19.4%	11.3%
CNBC	$0.34	$0.93	170.6%	$0.53	$0.70	30.7%	43.7%	13.6%
CNN	$0.49	$2.92	498.0%	$0.81	$1.98	144.1%	265.3%	7.2%
Comedy Central	$0.23	$0.66	187.5%	$0.61	$0.72	18.2%	43.2%	5.8%
Country Music TV	$0.18	$0.56	211.1%	$0.26	$0.29	10.8%	17.7%	0.2%
Court TV	$0.22	$0.85	276.1%	$0.35	$0.49	41.5%	63.9%	12.2%
Discovery Channel	$0.34	$1.47	339.6%	$0.59	$1.16	95.9%	182.0%	10.0%
Disney Channel	$0.77	$0.70	-8.9%	$0.68	$0.27	-59.6%	-59.6%	0.0%
E! Entertainment Television	$0.30	$0.48	62.0%	$0.41	$0.38	-7.6%	-15.8%	7.2%
ESPN	$2.44	$0.87	-64.5%	$3.80	$2.33	-38.6%	-75.9%	9.5%
ESPN 2	$0.33	$0.71	114.2%	$0.46	$0.48	3.9%	1.8%	7.7%
Food Network	$0.19	$0.85	352.9%	$0.49	$0.71	44.0%	122.1%	4.5%
Fox News Channel	$0.36	$1.83	411.8%	$0.70	$1.27	82.4%	171.8%	8.9%
Fox Sports Net	$1.56	$0.79	-49.3%	$1.51	$0.46	-69.4%	-77.4%	8.9%
FX	$0.36	$0.68	90.3%	$0.61	$0.58	-5.3%	-19.8%	10.2%
GSN	$0.19	$0.42	124.3%	$0.23	$0.12	-47.7%	-76.0%	20.7%
Golf Channel	$0.32	$0.14	-57.5%	$0.37	$0.10	-72.6%	-99.9%	14.9%
Hallmark Channel	$0.17	$0.63	272.5%	$0.33	$0.32	-3.7%	-28.6%	17.1%
HGTV	$0.25	$1.04	310.8%	$0.60	$0.82	38.4%	77.2%	15.2%
History Channel	$0.29	$2.29	699.5%	$0.53	$1.16	120.5%	237.0%	13.5%
Lifetime	$0.32	$0.85	166.8%	$0.81	$0.88	9.3%	-4.6%	16.7%
MSNBC	$0.26	$0.69	168.3%	$0.33	$0.31	-4.8%	-14.6%	16.1%
MTV	$0.37	$0.47	28.3%	$1.02	$0.93	-8.4%	-44.6%	8.6%
MTV2	$0.17	$0.54	223.0%	$0.19	$0.21	9.4%	12.4%	-0.5%
National Geographic Channel	$0.29	$0.65	120.9%	$0.34	$0.32	-5.1%	-6.2%	-1.2%
Nickelodeon	$0.48	$0.45	-7.5%	$1.38	$1.23	-10.5%	-61.8%	12.5%
Oxygen	$0.24	$0.09	-63.7%	$0.31	$0.16	-48.0%	-76.1%	16.5%
Syfy	$0.27	$0.70	160.0%	$0.55	$0.63	15.3%	18.3%	13.0%
SoapNet	$0.22	$0.44	98.8%	$0.24	$0.15	-37.9%	-47.0%	3.7%
Speed Channel	$0.27	$0.42	56.7%	$0.32	$0.18	-43.9%	-51.8%	-21.3%
Spike TV	$0.29	$0.60	106.7%	$0.54	$0.53	-1.1%	-8.6%	5.8%
TBS Superstation	$0.38	$0.88	132.0%	$0.89	$1.04	16.5%	33.1%	6.6%
The Weather Channel	$0.22	$0.60	174.4%	$0.34	$0.56	64.7%	102.4%	15.1%
TLC	$0.27	$0.83	205.9%	$0.42	$0.57	35.7%	55.5%	9.5%
TNT	$0.84	$0.93	11.1%	$1.35	$1.15	-15.2%	-33.6%	6.9%
Toon Disney	$0.21	$0.39	86.1%	$0.24	$0.10	-57.9%	-83.2%	17.7%
Travel Channel	$0.26	$0.45	69.7%	$0.32	$0.16	-50.5%	-74.9%	14.4%
TV Guide Channel	$0.16	$0.14	-16.2%	$0.24	$0.18	-24.3%	-49.4%	15.9%
TV Land	$0.21	$0.86	301.1%	$0.34	$0.53	57.0%	92.8%	11.9%
USA Network	$0.51	$0.84	65.0%	$1.13	$1.17	3.7%	-12.2%	14.1%
Versus	$0.25	$0.29	17.7%	$0.26	$0.13	-51.8%	-60.4%	-8.9%
VH1	$0.24	$0.44	80.8%	$0.55	$0.50	-9.7%	-27.3%	1.4%
WE: Women's Entertainment	$0.22	$0.32	46.1%	$0.26	$0.19	-28.5%	-39.8%	5.1%
Total	$18.22	$36.98	103.0%	$29.41	$30.16	2.6%	-3.7%	10.1%

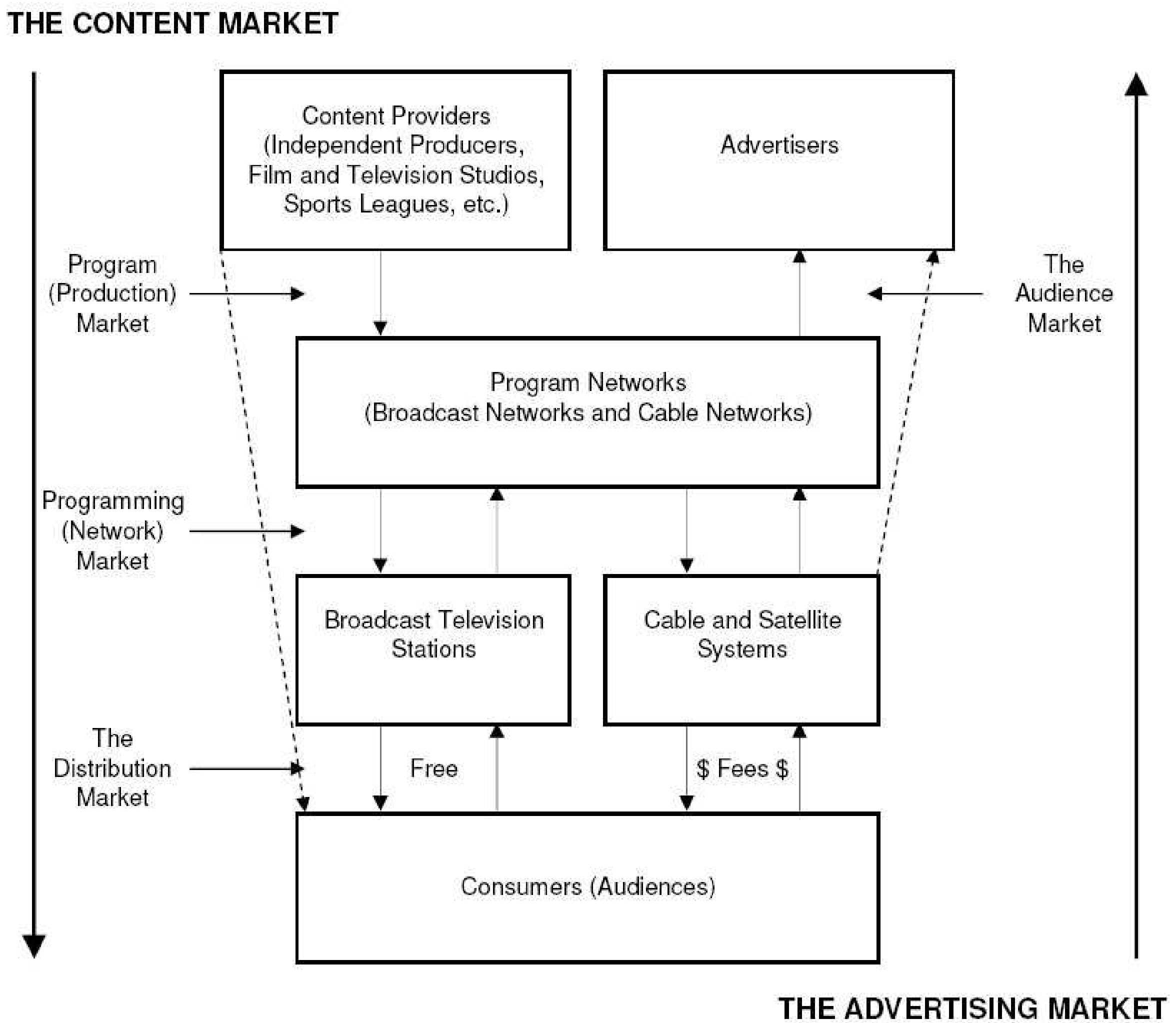
Notes: This table reports the results by channel of the input cost and profit consequences from our baseline, Full À La Carte (ALC), counterfactual with input cost renegotiation. As in Table 8, downstream prices are set at the renegotiated input costs; see footnote 40 for details. The first three columns report weighted averages (across distributors) of our estimated per-subscriber input costs under bundling and ALC equilibria (and their associated change). They are measured in 2000 dollars per subscriber per month. Distributors must pay the bundle input cost for all their subscribers in the bundling counterfactual, but pay the ALC input cost only for those that choose to subscribe under the ALC counterfactual. The remaining columns summarize the profit effects by channel. The fourth through seventh columns report the total (license fee plus advertising) profit effects, while the last two columns break out the percentage change for each of these components. Profits are measured in 2000 dollars per household per month.

Table 10: Alternative Counterfactual: Full ALC, Bundle-Sized Pricing, and Theme Tiers

		Levels			Percent Change		
	Bundling	Full ALC	Bundle Sized Pricing	Theme Tiers	Full ALC	Bundle Sized Pricing	Theme Tiers
Non-welfare Outcomes							
Cable & Sat Penetration	0.880	0.993	0.987	0.977	12.8%	12.1%	11.0%
Total Affiliate Fees	$18.22	$36.98	$34.44	$35.49	103.0%	89.1%	94.9%
Mean Consumer Expn	$27.63	$28.24	$28.60	$36.98	2.2%	3.5%	33.8%
Number Channels Received	42.8	19.3	17.0	34.7	-54.9%	-60.3%	-18.8%
Number Channels Watched	22.2	19.3	15.8	19.2	-12.8%	-28.7%	-13.4%
Welfare Outcomes							
Channel Profits							
Total License Fee Rev	$16.03	$15.44	$17.97	$25.26	-3.7%	12.0%	57.5%
Total Advertising Rev	$13.38	$14.73	$14.44	$13.95	10.1%	7.9%	4.3%
Total Channel Revenue	$29.41	$30.16	$32.40	$39.20	2.6%	10.2%	33.3%
Distributor Profits	$11.59	$12.81	$10.63	$11.72	10.4%	-8.3%	1.1%
Total Industry Profits	$41.00	$42.97	$43.03	$50.93	4.8%	5.0%	24.2%
Mean Consumers Surplus	$45.82	$45.91	$41.79	$35.73	0.2%	-8.8%	-22.0%
Total Surplus	$86.82	$88.88	$84.82	$86.66	2.4%	-2.3%	-0.2%

Notes: This table reports the results of alternative counterfactual simulations of various policy interventions on prices and welfare. The economic environment is as in Table 8. Columns one, two, and five report the counterfactual outcomes in bundling and full à la carte (ALC) environments as in Table 8. The remaining columns report counterfactual outcomes under Bundle-Sized Pricing and Theme Tiers. In the Bundle-Sized Pricing counterfactual, each downstream distributor competes by offering a full bundle of all the channels and a second bundle of fifteen channels, the identities of which may be chosen by each household. In the Theme Tier counterfactual, each downstream distributor competes by setting a fixed fee and offering 5 theme tiers from which the household can choose any combination. The theme tiers are Sports (ESPN, ESPN 2, Fox Sports Net, Golf Channel, Speed Channel, Versus), News (CNBC, CNN, Fox News Channel, MSNBC), Family and Education (ABC Family Channel, Animal Planet, Discovery Channel, Disney Channel, History Channel, National Geographic Channel, Nickelodeon, TLC, Toon Disney), Music and Lifestyle (Bravo, Country Music TV, E! Entertainment Television, Food Network, HGTV, Lifetime, MTV, MTV2, Oxygen, SoapNet, TV Guide Channel, VH1, WE: Women's Entertainment), and General (AMC, Arts & Entertainment, BET Networks, Cartoon Network, Comedy Central, Court TV, FX, GSN, Hallmark Channel, Syfy, Spike TV, TBS Superstation, The Weather Channel, TNT, Travel Channel, TV Land, USA Network). All counterfactuals allow for input-market renegotiation. Dollar values are 2000 dollars per U.S. television household per month.

Figure 5: Television Programming Industry



References

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

Andrews, D. and G. Soares, "Inference For Parameters Defined By Moment Inequalities Using Generalized Moment Selection," *Econometrica*, January 2010, *78* (1), 119–157.

Armstrong, M. and J. Vickers, "Competitive Price Discrimination," *RAND Journgal of Economics*, 2001, *32* (4), 579–605.

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

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

——, ——, and ——, "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 2004, *112* (1), 68–105.

Bertini, M. and L. Wathieu, "Attention Arousal Through Price Partitioning," *Marketing Science*, 2008, *27* (2), 236–246.

Byzalov, D., "Unbundling Cable Television: An Empirical Investigation," 2010. Working Paper, Temple University.

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

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

Chu, Chenghuan Sean, Phillip Leslie, and Alan Sorensen, "Nearly Optimal Pricing for Multiproduct Firms," 2010. forthcoming, *American Economic Review*.

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

——, "The Discriminatory Incentives to Bundle in the Cable Television Industry," *Quantitative Marketing and Economics*, 2008, *6* (1), 41–78.

Cremer, Jacques and Michael H. Riordan, "On Governing Multilateral Transactions with Bilateral Contracts," *The RAND Journal of Economics*, 1987, *18* (3), 436–451.

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

Fang, H. and P. Norman, "To Bundle or Not To Bundle," *Rand Journal of Economics*, 2006, *37* (4), 946–963.

FCC, "2001 Report on Cable Industry Prices," Technical Report, FCC 2002. Available at http://www.fcc.gov/mb/csrptpg.html.

———, "Report on the Packaging and Sale of Video Programming to the Public," Technical Report, FCC 2004. November 18, 2004. Available at http://www.fcc.gov/mb/csrptpg.html.

———, "Further Report on the Packaging and Sale of Video Programming to the Public," Technical Report, FCC 2006. February, 2006. Available at http://www.fcc.gov/mb/csrptpg.html.

———, "Thirteenth Annual Assessment on the Status of Competition in the Market for the Delivery of Video Programming (2006 Report)," Technical Report, Federal Communications Commission 2009. Available at http://www.fcc.gov/mb/csrptpg.html.

GAO, "Issues Related to Competition and Subscriber Rates in the Cable Television Industry," Technical Report, General Accounting Office October 2003. GAO-04-8.

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

Hart, Oliver and Jean Tirole, "Vertical Integration and Market Foreclosure," *Brookings Papers on Economic Activity. Microeconomics*, 1990, *1990*, 205–286.

Hausman, J., "Valuation of New Goods under Perfect and Imperfect Competition," in T. Bresnahan and R. Gordon, eds., *The Economics of New Goods*, University of Chicago Press, 1996.

———, G. Leonard, and D. Zona, "Competitive Analysis with Differentiated Products," *Annales d'Economie et de Statistique*, 1994, *34*, 159–180.

Ho, Katherine, "Insurer-Provider Networks in the Medical Care Market," *American Economic Review*, 2009, *99* (1), 393–430.

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

Kagan World Media, "Economics of Basic Cable Television Networks," Technical Report, Kagan World Media 2008.

Lee, Robin S., “Dynamic Demand Estimation in Platform and Two-Sided Markets,” 2010. Working Paper, NYU Stern.

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

Muthoo, Abhinay, *Bargaining Theory with Applications*, Cambridge: Cambridge University Press, 1999.

Nash, John F., “The Bargaining Problem,” *Econometrica*, 1950, *18* (2), 155–162.

Nevo, A., “Measuring Market Power in the Ready-To-Eat Cereal Industry,” *Econometrica*, 2001, *69* (2), 307–342.

____ **and M. Rosen**, “Identification with Imperfect Instruments,” 2008. Cemmap Working Paper CWP16/08.

Pakes, A., J. Porter, K. Ho, and J. Ishii, “Moment Inequalities and their Application,” 2007. Working Paper, Harvard University.

Petrin, A., “Quantifying the Benefits of New Products: The Case of the Minivan,” *Journal of Political Economy*, 2003, *110* (4), 705–729.

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

Rennhoff, A. and K. Serfes, “The Role of Upstream-Downstream Competition on Bundling Decisions: Should Regulators Force Firms to Unbundle?,” *Journal of Economics and Management Strategy*, 2009, *18* (2), 547–588.

Rennhoff, Adam D. and Konstantinos Serfes, “Estimating the Effects of a la Carte Pricing: The Case of Cable Television,” *SSRN eLibrary*, 2008.

Reuters, “U.S. Lawmaker Urges A La Carte Cable Channel Rates,” *Reuters News Service*, 2003. March 14, 2003.

Rochet, J.-C. and L. Stole, “Nonlinear Pricing with Random Participation,” *Review of Economic Studies*, 2002, *69* (1), 277–311.

Rubinstein, Ariel, “Perfect Equilibrium in a Bargaining Model,” *Econometrica*, 1982, *50* (1), 97–109.

Schmalensee, R., “Gaussian Demand and Commodity Bundling,” *Journal of Business*, 1984, *62*, S211–S230.

Segal, Ilya and Michael D. Whinston, “Robust Predictions for Bilateral Contracting with Externalities,” *Econometrica*, 2003, *71* (3), 757–791.

Shatz, Amy, "À la Carte Pricing May Cut Bills For Cable Customers, FCC Says," *Wall Street Journal*, 2006. February 10, 2006.

Stigler, G., "United States v. Loew's Inc.: A Note on Block Booking," in P. Kurland, ed., *The Supreme Court Review: 1963*, University of Chicago Press, 1963, pp. 152–157.

Yurukoglu, A., "Bundling and Vertical Relationships in Multichannel Television," 2009. mimeo, NYU Stern.

This Appendix Not for Publication

B Data Quality, Counterfactual Robustness, and Appendix Tables

B.1 Data Quality

Warren Factbook Data The Factbook data suffers from two weaknesses: persistent non-updating of entries and incomplete observations. When comparing yearly entries on an individual cable system in the Factbook, it is common to see that data does not change between two (and sometimes several) years. Given industry subscriber churn rates, channel introduction during the relevant time periods, and pricing behavior, we are certain that a lack of updating is the cause. Another common occurrence when analyzing the Factbook is that a cable system will have a bundle on offer, but no price and/or quantity is listed. Similarly, some observations are missing the number of homes the cable system passes. We try to estimate this figure when possible using census data on number of households. Sometimes this estimation is obviously unsuccessful, producing market shares well over one, for example. A third dimension of incomplete data in the Factbook deals with geographical market definition. In a few geographical markets, particularly dense metropolitan areas, there is more than one cable system. However, the Factbook does not specify on what portions of the market the cable systems overlap. We drop any observation for which there is a common community served with a distinct cable system, or if Factbook designates the system an overbuild. We present statistics on the extent of these two data quality issues below in Table 15. As can be seen there, the share of observations in a given year that are full and complete varies from 2% (in 2005) to 41% (in 1997).

While we worry in general about the quality of the Factbook data and its suitability for extrapolation to cable systems as a whole, we don't think it poses a serious econometric issue. In particular, we don't think unobservable characteristics of cable systems that impact whether an entry in the Factbook is up-to-date are likely to be correlated with the demand they face and/or their pricing behavior.

Satellite Data As noted in the text, we only observe market shares for the aggregate of bundles offered by both satellite providers at the DMA level. To accommodate this data limitation, we make the following two assumptions in our modeling approach. First, we assume the only satellite bundle in the DMA is the DirecTV total choice bundle (the most popular satellite bundle offered by either provider). Second, within a DMA, we assume the unobservable quality measure of this bundle does not vary across systems.

Ratings Data Nielsen is the dominant provider of television ratings. It has a large staff dedicated to data quality, statistical integrity, and metering technology. Our data comes from Set Meters which measure electronically to what channel the television is tuned throughout the day. This data is then linked with which programs aired on the relevant channels. We therefore have considerable confidence in the quality of the ratings data.[47]

B.2 Counterfactual Robustness

Our goal is to accurately measure the welfare effects of à la carte pricing in multichannel television markets. As such, it is important to have confidence that our qualitative results are robust and not sensitive to particular assumptions underlying the counterfactual exercises. In this sub-section we consider the robustness of our results to alternative assumptions on downstream markups, demand, and bargaining in our counterfactual exercises.

Due to the computational cost of estimating the full model, all of these robustness exercises are undertaken for the counterfactual analysis only.[48] The method used to appropriately conduct the counterfactual under each alternative assumption varied; the specifics for each are described below.

We evaluated the robustness of our results in the following dimensions:

Downstream Markups As described in footnote 40 in the text, for computational reasons we assume that downstream channel markups are zero in our counterfactual analysis and that distributors instead earn profit on the fixed fees that they charge. In this robustness exercise, we allow downstream margins to be 10% instead of zero. This is at the upper end of the range we were finding when we tried to flexibly solve for them in the counterfactual equilibrium.

Demand: Marginal Distributions One of the critical assumptions underlying our demand model is the shape of households' distribution of preferences (WTP) for the individual channels that constitute existing service bundles. As discussed in Section 5 and motivated by our individual-level data as shown in Figure 3, we assume that the marginal distribution of unobserved tastes for each channel is a mixture of a mass point at zero and an exponential distribution whose (single) mean and variance parameter we estimate for each channel. To evaluate the robustness of this assumption, we conducted our counterfactual analysis under two alternative families of marginal distributions: the Rayleigh Distribution and the Log-Normal Distribution. The Rayleigh distribution is also a single-parameter family, but, relative to the exponential, it has a slightly smaller coefficient of variation (COV), a non-zero mode, and smaller skewness and kurtosis. It looks a bit like a log-normal, but with a thinner right tail than both it and the exponential. The Log-Normal distribution is a two-parameter family which, for mean and variance comparable to those we find for individual channels using our exponential distribution, also has a non-zero mode and larger skewness and kurtosis. With these

[47] That being said, it is not without its critics. Nielsen data has been criticized both for not accurately capturing the whole television universe, for example out-of-home viewing, and for sample sizes too small to accurately measure the viewing of niche programming.

[48] For example, estimating the full demand model under alternative assumptions for marginal distributions would take several weeks for each assumption considered.

choices, we are effectively allowing tastes to (1) have more mass nearer the center of the distribution and (2) relatively thinner or thicker tails than an exponential.

To evaluate the robustness of our distributional assumption on the marginals, we maintain the assumption of the zero mass point,[49] but calibrate the parameters of the Rayleigh or Log-Normal for each channel to match as closely as possible the implied mean and variance of the estimated WTP for that channel. We then re-estimated our Full ALC counterfactual using these implied marginal distributions and the input costs implied by renegotiation under the exponential distribution.[50]

Demand: Correlations One of the primary motivations for bundling identified in the theoretical literature is the degree of correlation in tastes for bundle components. We allow for correlation from both demographic differences in tastes as well as correlation in unobserved tastes. We evaluate the robustness of our findings to these correlations by conducting our Full ALC counterfactual eliminating unobserved correlations.[51] To do so, we set all off-diagonal elements of the covariance structure of our estimated $G()$ distribution to zero. For the same reasons as for the marginal distribution calculations above, we do so at the renegotiated input costs implied by the full (with correlation) model.

Bargaining: Halve/Double Input Costs A key element of this paper is our ability to estimate bargaining parameters and predict renegotiated input costs in an ALC environment. It is possible, however, that true bargaining outcomes would differ from our predictions. To get a sense of how important this might be, we evaluate our Full ALC counterfactual under two different assumptions: that estimated input costs are either half or double our estimated renegotiated values.

When a Channel is Watched Less but Valued More (Monte Carlo) In our model, we assume that channels that are viewed more are valued more by households. It is possible, however, that minutes of different types of programming provide different utility profiles. For example, some programming (e.g. sports programming) may provide more value to households even if it is watched for fewer minutes than other programming. We explore the consequences of this possibility using monte carlo simulation.

[49] It is an important factor allowing us to accurately predict the number of channels watched by households when offered a bundle of channels.

[50] Using the renegotiation input costs under our exponential assumption was also necessary due to the high computational costs of calculating renegotiation equilibria. Overall mean WTP for the bundle under the alternative distributions differed slightly from that coming out of the exponential. To ensure comparability across the counterfactuals, we allocated this mean WTP difference to CS and/or Profit at the same proportion as that implied by the counterfactual for that distributional family.

[51] It is more complicated to eliminate correlations due to demographics as they influence both the mean and variance-covariance matrix of tastes for channels. Because demographics explained only 5% of the variation in mean tastes, we decided to simply eliminate correlation due to the unobserved component.

B.2.1 Markup, Demand, and Bargaining Robustness

Table 16 at the end of this Appendix reports the results of each of our robustness exercises except the monte carlo. For each different assumption considered, we report the percent change in consumer surplus, industry profit, and total surplus. The first row replicates these values for our baseline, Full À La Carte counterfactual.

Assuming the larger 10% markup downstream reduces all of consumer, firm, and total welfare relative to the Full ALC baseline. This is due to the standard consequences of double marginalization: prices are higher (reducing consumer welfare), but total industry profits and total surplus decline. Each of the predicted changes is small relative to the Full ALC baseline and therefore yield qualitatively similar conclusions.

Changes in demand assumptions have slightly larger effects. Assuming preferences are distributed according to a Rayleigh (Log-normal) distribution yields lower (unchanged) consumer surplus and lower (higher) industry profits. These suggest firms are profiting from high-valued consumers in the tails of the taste distributions under ALC. Eliminating correlations reduces consumer surplus and increases profit, suggesting the overall pattern of correlations in the estimated preferences is positive. Similar to the effect of correlation on demand for bundles, eliminating this positive correlation reduces the heterogeneity in household WTP for their preferred channels under ALC, increasing firm profits and reducing consumer surplus. None of these effects, however, materially change the conclusions about the welfare effects of ALC.

By contrast, alternative bargaining assumptions have substantial effects on our estimated welfare changes. Recall the total increase in input costs under our baseline counterfactual was an estimated 103.0%. If we halve those, we find a substantially different picture: consumer welfare increases considerably (+18.5%), industry profits fall (-10.1%), and total surplus increases. These effects are qualitatively similar to that which we found when evaluating the welfare effects while keeping input costs at their level in a bundling equilibrium: it is the sharp rise in input costs (and prices) that prevents a significant increase in consumer welfare under ALC. Doubling our estimated renegotiation input costs would, not surprisingly, be even worse for consumers, reducing consumer surplus by an estimated 27.6%. Industry profits rise significantly in this setting and total welfare falls.

Across these robustness exercises, only the changes in bargaining outcomes have a meaningful impact on the magnitude of our estimated welfare effects. How then should one interpret them? *If* our assumptions on renegotiated input costs under à la carte are incorrect, we conclude that because a doubling of input costs increases industry profits, that makes it the more likely of the two deviations. If so, prospects are even worse for consumer and total welfare than in our baseline results presented in the body of the text. Like our baseline, these results also do not take into account any additional implementation or marketing costs that might arise in an à la carte environment. We therefore conclude that our qualitative conclusions about à la carte are robust: in the absence of input costs changes, it would likely improve consumer welfare, but in their presence, consumers are likely better off with existing bundles.

B.2.2 When a Channel is Watched Less but Valued More (Monte Carlo)

To allow for the possibility that a channel that is watched less than another is nonetheless valued more, we begin by specifying a richer model for consumer utility than that used in the paper. We consider the case of three goods: two television channels, $c = \{1, 2\}$, and an outside good, denoted $c = 3$. Let c index channels and assume all households face a bundle with both channels. We assume the utility to household i from spending their time watching television and doing non-television activities has the following form:

$$v_i(t_i) = \sum_{c \in \{1,2,3\}} \frac{\gamma_{ic}}{1 - \nu_{ic}} (1 + t_{ic})^{1-\nu_{ic}} \tag{8}$$

where t_i is a vector with components t_{ic} which denote the number of hours household i watches channel c. As in the Cobb-Douglas specification in Equation (1) in the text, γ_{ic} is a parameter representing i's tastes for channel c. The novelty in this specification is ν_{ic}. ν_{ic} governs the shape of marginal utility household i obtains from watching channel i. Marginal utility in this specification is $\frac{\partial v_i}{\partial t_{ic}} = \frac{\gamma_{ic}}{(1+t_{ic})^{\nu_{ic}}}$. As $\nu_{ic} \to 1$, $\forall c$, this functional form converges to the Cobb-Douglas specification, with relatively steep decreases in marginal utility across minutes. As $\nu_{ic} \to 0$, marginal utility converges to a constant across-minute value, γ_{ic}.

Interesting patterns can result from this specification when households have high values of γ_{ic} and ν_{ic} for some channels and low values for others. Figure 7 provides an example. This figure presents graphically the optimal decision-making for a household with preference parameters, $\gamma_c = [2.5 \;\; 6.0 \;\; 2.9]$ and $\nu_c = [0.2 \;\; 0.9 \;\; 0.2]$. For convenience, we omit the i subscript. Let $v_c = \frac{\gamma_c}{1-\nu_c}(1 + t_c)^{(1-\nu_c)}$ be defined as the contribution channel c makes to the household's utility. The left-hand panel presents household i's utility for various values of time spent watching channel 1, given the optimal time spent not watching TV (which, for these parameters, is $t_3 = 14.3$ hours).[52] The increasing, dashed line and the decreasing, dotted line plot the utility channel 1 and 2 contribute to total utility, given by the solid line at the top of the left panel. Utility from channel 1 (channel 2) increases (decreases) with time spent watching channel 1. The optimal time spent watching channel 1 (channel 2) is 6.6 (3.1) hours. The optimal t_1^* is denoted in both panels with a vertical dashed line. In the right-hand panel is shown that this optimum is obtained at the point where the marginal utility of an additional minute watching channel 1 (again given by the dashed line) equals the marginal utility of an additional minute watching channel 2 (again given by the dotted line).

What is different about channels 1 and 2 is the shape of marginal utility for minutes provided by each. Channel 1 has relatively low values of γ and ν and consequently has a relatively flat marginal utility profile in the right-hand panel. Channel 2 has relatively high values of γ and ν and a relatively steep marginal utility profile (when read from the right axis). The consequence of these shapes is that channel 2 contributes relatively more to the household's utility, *despite being watched for fewer minutes*. This is captured in the figure by u_c^*.[53] For these parameter values, channel 2 is watched less than half as much as channel 1 (3.1

[52] The figure is constructed so that time spent watching channel 2 is given by the distance from the right on the horizontal axis.

[53] u_c^* measures the maximum utility from watching all the channels less the maximum utility when not watching channel c, e.g. $u_1^* = v(t_1^*, t_2^*, t_3^*) - v(0, t_2^{**}, t_3^{**})$ where t_{-c}^{**} are the optimal times spent watching channels other than channel c.

hours versus 6.6 hours), but the household would be willing to pay twice as much for it (4.0 utils versus 2.0 utils).

Monte Carlo Simulation We explore the consequences of preferences like those in Figure 7 using monte carlo simulation. We first generate data from a true distribution modeled on that described above. We then estimate parameters based on a Cobb-Douglas utility model like that estimated in our paper. Finally, we compare the difference in aggregate market outcomes (market shares, prices, welfare measures) in the true data compared to what we estimate. We describe each of these stages in turn.

Data Generated from True Preferences We generate data from a single market with a distribution of households with preferences, $\theta_i \equiv (\gamma_i' \;\; \nu_i')'$, whose means are similar to those in Figure 7. Based on their preference draws for channels, individual households decide how much they would watch each channel if they purchased the bundle, t_i^*, and compare their utility from that viewing, $v_i^*(t_i^*)$, to the bundle price plus a random error distributed as a type 1 extreme value.

$$u_i \;\; = \;\; v_i^*(t_i^*(\theta_i)) - p + \epsilon_i \tag{9}$$

The predicted market shares for the bundle is then just

$$s_b(p,\theta) \;\; = \;\; \int \frac{exp((v_i^*(t_i^*(\theta_i))) - p)dF(i)}{1 + exp((v_i^*(t_i^*(\theta_i))) - p)} \tag{10}$$

where θ is a vector with typical element, θ_i and $dF(i)$ is the true distribution of θ_i in our population. We approximate this integral using simulation with 150 households. We assume a vertically integrated monopoly programmer/distributor sells a bundle of two television channels at its short-run profit-maximizing price facing zero marginal costs, implying an optimal bundle price, p_b. This implies an equilibrium market share, $s_b(p_b, \theta)$, as well as mean consumer surplus, $CS_b(\theta)$, profit, $\Pi_b(\theta)$, and total surplus, $TS_b(\theta)$.

Based on these true preferences, we also simulate outcomes when the monopolist sells products à la carte (ALC). In this case, a household's choice set now has four options: channel 1 alone, channel 2 alone, a bundle of both channels, or no television. The logic of preferences, market shares, optimal prices, and welfare measures follows analogously to the bundle case, implying true ALC prices, $p_{alc} = \{p_{alc,1}, p_{alc,2}\}$[54], market shares, $s_{alc,c}, c \in \{1, 2, b, 3\}$, consumer surplus, CS_{alc}, profit, Π_{alc}, and total surplus, TS_{alc}.

Estimation Model To assess the biases from having less-watched channels be more valuable, we outline our model used for estimation. It is similar to the true model, except that we assume a straight Cobb-Douglas utility specification analogous to that used in our paper rather than the richer utility function used to generate

[54] With $p_{alc,b} = p_{alc,1} + p_{alc,2}$.

the data.

$$v_i(t_i) = \sum_{c \in \{1,2,3\}} \gamma_{ic} \log(1 + t_{ic}) \tag{11}$$

As in the paper, we assume γ_{ic} is drawn from a distribution with known parameters, ϕ. In practice, we will assume $\phi \equiv [E\gamma_{ic} \;\; \sigma_{\gamma_{ic}}]$, i.e. the mean and standard deviation of γ_{ic} for $c = \{1, 2, 3\}$, with $\gamma_{i3} = 10, \forall i$ normalized to set the scale of utility. Based on these assumed preferences, households decide how much they would watch each channel if they purchased the bundle, t_i^*, and compare their utility from that viewing, $v_i^*(t_i^*)$, to the bundle price plus a random error distributed as a type 1 extreme value.

$$u_i \;\; = \;\; v_i^*(t_i^*(\phi_i)) - \alpha p + \epsilon_i \tag{12}$$

where α is an (estimated) parameter measuring marginal utility of income ($\alpha = 1$ in the true data generating process). Let $\delta = (\phi' \;\; \alpha)'$ define the vector of (six) parameters to be estimated.

The predicted market shares for an estimated bundle is then just

$$s_b(p, \delta) \;\; = \;\; \int \frac{exp((v_i^*(t_i^*(\phi_i))) - \alpha p) dF_\phi(i)}{1 + exp((v_i^*(t_i^*(\phi_i))) - \alpha p)} \tag{13}$$

where $dF_\phi(i)$ is the distribution of γ_i in our estimation model. As above, we approximate this integral using simulation with 150 households. We continue to assume a vertically integrated monopoly programmer/distributor that sells a bundle of two television channels at its short-run profit-maximizing price facing zero marginal costs, implying an optimal estimation model bundle price, p_b^{cd}, where cd stands for Cobb-Douglas, our estimation model. This implies an equilibrium market share, $s_b^{cd}(p_b^{cd}, \delta)$, as well as mean consumer surplus, $CS_b^{cd}(\delta)$, profit, $\Pi_b^{cd}(\delta)$, and total surplus, $TS_b^{cd}(\delta)$.

Once we have estimated parameters for this model, we also simulate outcomes from the estimation model when the monopolist sells products à la carte (ALC). As earlier, a household's choice set now has four options: channel 1 alone, channel 2 alone, a bundle of both channels, or no television. The logic of preferences, market shares, optimal prices, and welfare measures follows analogously for the bundle case, implying estimation model ALC prices, $p_{alc}^{cd} = \{p_{alc,1}^{cd}, p_{alc,2}^{cd}\}$, market shares, $s_{alc,c}^{cd}, c \in \{1, 2, b, 3\}$, consumer surplus, CS_{alc}^{cd}, profit, Π_{alc}^{cd}, and total surplus, TS_{alc}^{cd}.

Estimation For estimation, we must generate the true data, calculate moments for that true data, and compare those moments to moments generated by our estimation model in order to estimate the parameters of that estimation model.

The true data were generated with six free utility parameters.[55] The eight moments we use in estimation are the mean and standard deviation of average viewing time for the three channels and the mean and variance of the bundle market share.[56]

[55] $\gamma_{i1} \sim N(2.5, 0.1^2)$, $\gamma_{i2} \sim N(6, 0.5^2)$, $\gamma_{i3} \equiv 3, \forall i$, $\nu_{i1} \sim U[.15, .25]$, $\nu_{i2} \sim U[.85, .95]$, $\nu_{i3} \sim U[.15, .25]$.

[56] The mean and standard deviation of viewing times depends on whether or not households choose to purchase the bundle. Thus our estimate of the expected true viewing time for channel 1 is $Et_1 = \sum_i [t_{1i} * s_{ib} + 0 * (1 - s_{ib})]$, where $s_{ib} = \frac{exp((v_i^*(t_i^*(\theta_i))) - p)}{1 + exp((v_i^*(t_i^*(\theta_i))) - p)}$.

The estimation model also predicts outcomes with six free utility parameters.[57] We calculate the same eight moments from the estimation model to compare with the moments from the true data. In estimation, we weight the difference between each of the "true" and predicted moments equally.

Results: Channels Table 17 presents a summary of the results of the monte carlo exercise. In the first group of columns it presents various outcomes for bundled and à la carte market structures for both the true data and the estimates based on the Cobb Douglas utility function like that used in the analysis in the paper. Reported there are the mean (across 40 replications) ratios of outcomes for channel 1 relative to the same outcomes for channel 2. For example, the first cell in the table reports that, based on the true parameters, the expected viewing of channel 1 is 1.98 times the expected viewing of channel 2 when both are offered in a bundle. The cell adjacent to that reports that, based on the *estimated* parameters, the same ratio of expected viewing of channel 1 to channel 2 in a bundle is 2.12. Below these values are the standard deviation of these values across the 40 replications in our monte carlo study. This first cell suggests the estimated model predicts ratios of viewing times similar to those generated by the true parameters.

The balance of the ratios in the rest of the cells in the first set of columns yield dramatically different conclusions. The richer viewership model induces extreme biases in expected WTP for bundles and prices and shares of each channel in an ALC environment. For example, because channel 2 has high utility at low minutes (and channel 1 the opposite), the expected WTP for channel 1 in the true data is less than half (0.46) that for channel 2. By contrast, estimating a Cobb-Douglas model implies, like estimates of expected viewing time, that expected WTP for channel 1 is much greater (3.42 times) than that for channel 2. We similarly mis-estimate prices and market shares for the channels in an ALC world: in each case the true model implies higher values for channel 2 than channel 1, but the estimates imply the opposite.

Results: Welfare The key question for the conclusions in this paper is whether these striking biases at the channel level translate into mis-estimates of the welfare effects of ALC when aggregating *across* channels. The results in the second set of columns demonstrate that this *isn't* the case. Reported there is the mean (across replications) percentage difference in aggregate consumer surplus, profit, and total surplus moving from a bundling to an ALC environment.[58] In the absence of input cost changes, we anticipate that consumers benefit from ALC and the table shows that to be the case: aggregate consumer surplus increases an expected 15.7% for true preferences and by 11.6% in our estimated data. While these are different (by 4.1%), this difference isn't statistically significant. *Thus, while we badly mis-estimate outcomes for individual channel outcomes, aggregating across channels causes these errors to cancel out and yields no significant difference in estimated consumer surplus changes.* Similarly insignificant differences arise for profits and total surplus.

How is it possible to be so wrong for individual channels and not do so badly on average? The answer lies in

[57] $\gamma_{i1} \sim N(\gamma_1, \sigma_1^2)$, $\gamma_{i2} \sim N(\gamma_2, \sigma_2^2)$, $\gamma_{i3} \sim N(10, \sigma_3^2)$, and α.

[58] e.g., the first cell in the second group of columns calculates $\frac{CS_alc(\theta) - CS_b(\theta)}{CS_b(\theta)}$ associated with households' true preferences, θ.

the data: bundle purchase data like that in the monte carlo (and our paper) must at the end of the day equate total utility from viewing (based on households' total minutes of viewing) with the price of the bundle in a way to match market shares for bundles. Thus in the monte carlo (and the paper), we estimate something like an average-across-channels utility from minutes of viewing. If households value early minutes more highly for some channels (e.g. sports channels), we will tend to underestimate the utility (and prices and market shares) arising from those channels and overestimate these values for channels that have relatively flat marginal utility from minutes. But we won't be nearly as badly wrong about across-channel averages.

Figure 6: High and Low Rating DMA's for Black Entertainment Television

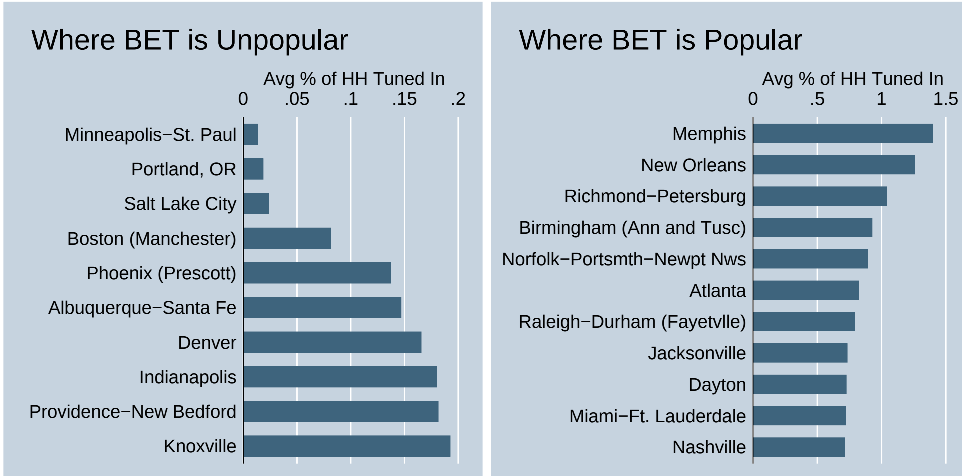


Table 11: Correlation in the Ratings Data

Network	Cartoon Network	A&E	Turner Classic Movies	Discovery Channel	ESPN	ESPN2	ESPN Classic	ESPN News
Cartoon Network	1							
A&E	-0.14	1						
TCM	-0.29	0.09	1					
Discovery	0.18	0.28	-0.33	1				
ESPN	0.14	0.01	0.07	-0.08	1			
ESPN2	0.11	0.16	0.10	0.08	0.54	1		
ESPN Classic	0.30	-0.10	0.16	-0.17	0.16	0.15	1	
ESPNews	0.35	-0.16	0.06	-0.09	0.26	0.20	0.39	1

Table 12: Sample Statistics, Other Estimation Data

Variable	NObs	Mean	SDev	Min	Max
Channel Dummies	See Tables in Paper				
Demographics					
Urban	56	0.61	0.22	0.14	0.99
Family	56	0.68	0.03	0.59	0.77
Household Income	56	$0.48	$0.07	$0.38	$0.75
Black	56	0.10	0.09	0.00	0.34
Hispanic	56	0.12	0.11	0.02	0.54
Asian	56	0.02	0.03	0.00	0.19
College Degree or Greater	56	0.18	0.06	0.09	0.36
Age	56	0.37	0.02	0.33	0.42

Table 13: Estimated Price Elasticities, B+EB+DB Markets

Price Elasticity of	wrt	Mean	Std. Dev.
Basic	Outside Good	0.15	0.27
	Basic	-4.12	2.25
	Expanded Basic	2.04	2.53
	Digital Basic	0.52	1.10
	Satellite	0.54	0.98
Expanded Basic	Outside Good	0.50	2.98
	Basic	0.16	0.51
	Expanded Basic	-6.34	2.99
	Digital Basic	2.12	2.64
	Satellite	1.52	1.47
Digital Basic	Outside Good	0.09	0.30
	Basic	0.09	0.78
	Expanded Basic	5.79	2.96
	Digital Basic	-13.11	4.10
	Satellite	2.56	1.89
Satellite	Outside Good	0.07	0.20
	Basic	0.05	0.41
	Expanded Basic	2.63	2.85
	Digital Basic	2.08	2.47
	Satellite	-5.35	3.44

Notes: B+EB+DB Markets are those offering Basic, Expanded Basic, and Digital Basic cable service.

Table 14: Carriage of Time Warner Channels by Distributor 2004-2007.

	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

Notes: 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.

Table 15: Data Quality of Factbook

Year	*Variable*	*Number of Bundles*	*Fraction of Bundles*
1997	Total Bundles	15,205	100.0%
	Full Information	10,740	71.0%
	Updated	9,264	61.0%
	Full Information and Updated	6,165	41.0%
1998	Total Bundles	15,743	100.0%
	Full Information	10,872	69.0%
	Updated	4,714	30.0%
	Full Information and Updated	3,461	22.0%
1999	Total Bundles	15,497	100.0%
	Full Information	10,444	67.0%
	Updated	5,663	37.0%
	Full Information and Updated	3,595	23.0%
2000	Total Bundles	15,453	100.0%
	Full Information	10,312	67.0%
	Updated	3,358	22.0%
	Full Information and Updated	2,478	16.0%
2001	Total Bundles	15,391	100.0%
	Full Information	9,793	64.0%
	Updated	4,173	27.0%
	Full Information and Updated	2,663	17.0%
2002	Total Bundles	15,287	100.0%
	Full Information	7,776	51.0%
	Updated	5,086	33.0%
	Full Information and Updated	1,484	10.0%
2003	Total Bundles	15,365	100.0%
	Full Information	8,370	54.0%
	Updated	9,744	63.0%
	Full Information and Updated	4,750	31.0%
2004	Total Bundles	15,145	100.0%
	Full Information	7,137	47.0%
	Updated	8,175	54.0%
	Full Information and Updated	3,556	23.0%
2005	Total Bundles	15,001	100.0%
	Full Information	7,009	47.0%
	Updated	846	6.0%
	Full Information and Updated	327	2.0%
2006	Total Bundles	14,653	100.0%
	Full Information	4,577	31.0%
	Updated	8,141	56.0%
	Full Information and Updated	2,303	16.0%
2007	Total Bundles	13,879	100.0%
	Full Information	4,070	29.0%
	Updated	3,135	23.0%
	Full Information and Updated	711	5.0%
1997-2007	Total Bundles	166,619	100.0%
	Full Information	91,100	55.0%
	Updated	62,299	37.0%
	Full Information and Updated	31,493	19.0%

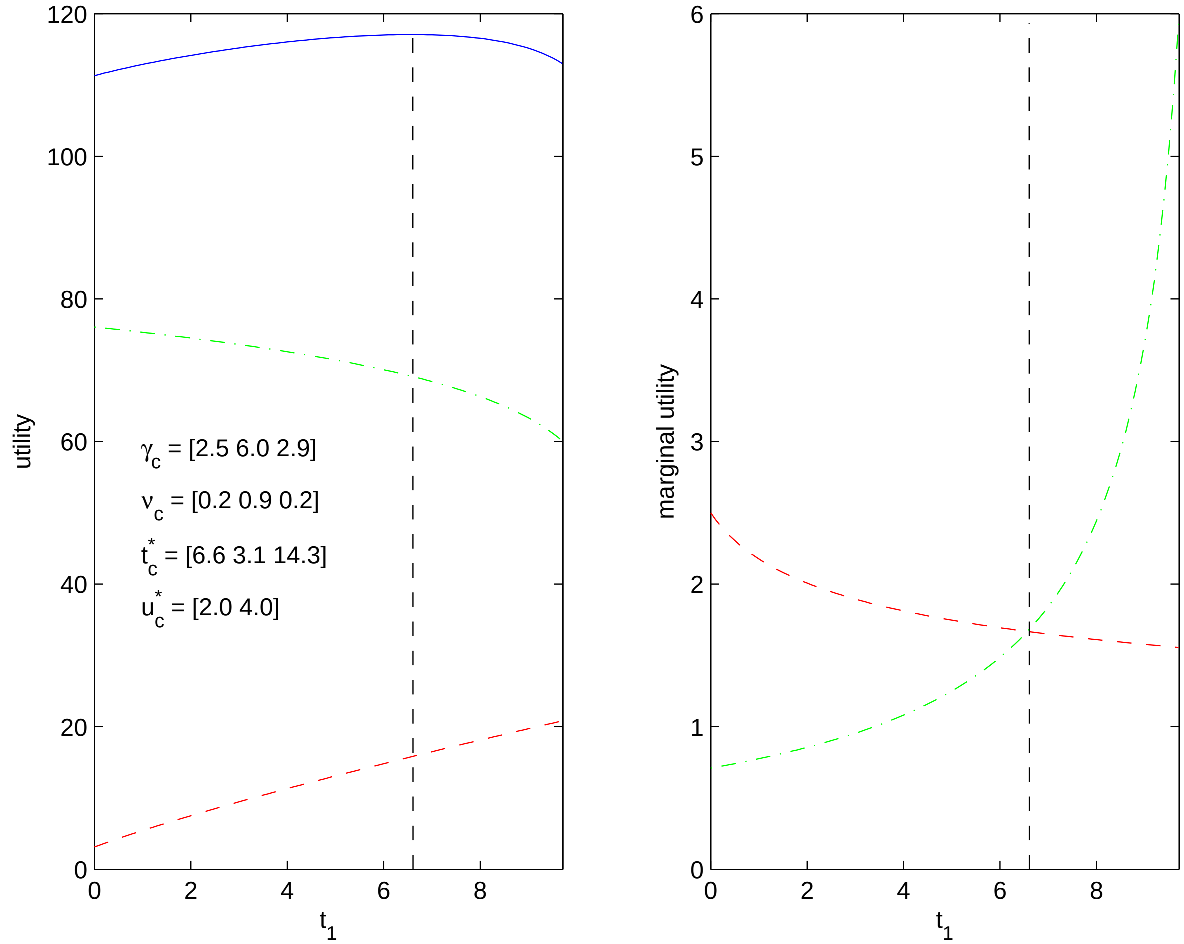
Notes:

Table 16: Robustness of Counterfactual Results

	% Change Consumer Surplus	% Change Industry Profit	% Change Total Surplus
Baseline Counterfactual			
Full À La Carte	0.2%	4.8%	2.4%
Alternative Distributor Markup			
10% Distributor Markup	-1.6%	2.5%	0.3%
Alternative Demand Assumptions			
Marginal Distributions: Rayleigh	-5.4%	2.4%	-1.7%
Marginal Distributions: Log-Normal	0.0%	12.8%	6.0%
Joint Distribution: No Correlation	-4.2%	8.6%	1.8%
Alternative Bargaining Assumptions			
Halve Input Costs	18.5%	-10.1%	5.0%
Double Input Costs	-27.6%	18.6%	-5.8%

Notes: This table reports the percentage change in consumer surplus, industry profits, and total surplus estimated under our baseline Full À La Carte counterfactual and under alternative assumptions about demand, bargaining conditions, downstream distributor markups, and/or exit in the counterfactual. All counterfactuals rely on parameter estimates from the baseline specification suitably adapted for the specific robustness test - see text for details. Alternative demand assumptions are evaluated at the renegotiated input costs from the baseline demand specification. The baseline counterfactual is as described in Table 10. See Appendix B.2 for a description of the specific alternative assumptions considered in the table.

Figure 7: When a Channel is Watched Less but Valued More



Notes: This figure demonstrates when a channel can be watched less than another but nonetheless be valued more. The data are generated from the alternative viewership model given by equation (8) in Appendix B.2 with values for γ_c and ν_c for each of three channels, c, given in the figure above. Optimal time viewing each channel is given by t^*_c. Willingness to pay for each channel, as defined in Appendix B.2, is given by u^*_c. The left panel reports the utility from channel 1 (dashed line), channel 2 (dotted line) and overall (solid line) as a function of time spent watching channel 1, t_1, given the optimal amount of time spent watching channel 3, $t^*_3 = 14.3$. The figure is constructed so that time spent watching channel 2 can be read from the right on the horizontal axis. The right panel reports the marginal utility for channels 1 (dashed) and 2 (dotted). The optimal viewing time is given in the right panel that equates the marginal utility for minutes for the two channels. This is at $t^*_1 = 6.6$ and $t^*_2 = 3.3$. Despite channel 2 being watched less, the household is willing to pay $u^*_2 = 4.0$ utils for access to it, double what it is willing to pay for channel 1 ($u^*_1 = 2.0$).

Table 17: When a Channel is Watched Less but Valued More: Monte Carlo Results

	True Preferences	Estimated Preferences		True Preferences	Estimated Preferences	Difference
	Mean (StdDev)	Mean (StdDev)		Mean (StdDev)	Mean (StdDev)	Mean (StdDev)
Ratio: $\text{Outcome}_1/\text{Outcome}_2$			% Difference, ALC - Bun			
Expected Viewing	1.98	2.12	Consumer Surplus	15.7%	11.6%	4.1%
Bundling	(0.05)	(0.12)		(0.7%)	(3.4%)	(3.5%)
Expected WTP	0.46	3.42	Profit	-11.4%	-5.9%	-5.5%
Bundling	(0.02)	(0.27)		(0.4%)	(3.8%)	(3.7%)
Prices	0.65	1.94	Total Surplus	-2.8%	-0.9%	-1.9%
A La Carte	(0.01)	(0.11)		(0.2%)	(1.4%)	(1.3%)
Prices	0.45	1.95				
A La Carte	(0.01)	(0.37)				

Notes: This table reports the results of a monte carlo simulation exercise to demonstrate the consequences when a channel is watched less by a household but is nonetheless valued more. As in Figure 7 above, channel 2 is watched less but valued more than channel 1. The first group of columns reports the mean (across 40 monte carlo replications) ratio of outcomes for channel 1 relative to the same outcomes for channel 2 for both true preferences given by equation (8) in Appendix B.2 as well as for estimates based on Cobb-Douglas utility like that estimated in the body of the paper. The second group of columns reports the mean percentage difference in aggregate welfare from a bundling to an à la carte environment for both true and estimated preferences as well as their difference. For all cells, the standard deviation across monte carlo replications is reported below the mean.