Quant Market Econ (2008) 6:41–78
DOI 10.1007/s11129-007-9031-7

The discriminatory incentives to bundle in the cable television industry

Gregory S. Crawford

Received: 19 June 2006 / Accepted: 14 June 2007 /
Published online: 26 September 2007

Abstract An influential theoretical literature supports a discriminatory explanation for product bundling: it reduces consumer heterogeneity, extracting surplus in a manner similar to second-degree price discrimination. This paper tests this theory and quantifies its importance in the cable television industry. The results provide qualified support for the theory. While bundling of general-interest cable networks is estimated to have no discriminatory effect, bundling an average top-15 *special-interest* cable network significantly increases the estimated elasticity of cable demand. Calibrating these results to a simple model of bundle demand with normally distributed tastes suggests that such bundling yields a heterogeneity reduction equal to a 4.7% increase in firm profits (and 4.0% reduction in consumers surplus). The results are robust to alternative explanations for bundling.

Keywords Bundling · Price discrimination · Cable television

JEL Classification L12 · M31 · L96 · L40 · L50 · C31

We are grateful to the editor and two anonymous referees for their detailed comments on the paper. We would also like to thank Cathleen McHugh for her assistance inputting the data, Mike Riordan, Joe Harrington, Matt Shum, Steve Coate, Roger Noll, Bruce Owen, V. Kerry Smith, Mark Coppejans, Frank Wolak, Phillip Leslie, and seminar participants at Cornell University and the 1999 IDEI/NBER Econometrics of Price and Product Competition conference for helpful comments.

G. S. Crawford (✉)
Department of Economics, University of Arizona,
Tucson, AZ 85721-0108, USA
e-mail: crawford@eller.arizona.edu

1 Introduction

Bundling is a customary feature of contemporary product markets. Telecommunications firms bundle local, long-distance, and mobile telephone services, banks bundle checking, credit, and investment services, and hospitals bundle an array of medical services. Even markets where consumers exhibit considerable discretion in choice, as for grocery stores or computer hardware, firms compete in broader dimensions of product breadth, specialty services, and convenience. Following Lancaster (1971), nearly all goods embody a bundle of attributes or characteristics.

Despite its prevalence, the microeconomic implications of product bundling are unclear. A variety of industries emphasize the benefits of bundling in simplifying consumer choice (as in telecommunications and financial services) or reducing costs from consolidated production of complementary products (as in health care and manufacturing). In either case, bundling promotes efficiency by reducing consumer search costs, reducing product or marketing costs, or both. More recently, focus has centered on bundling to extend market power or deter entry (e.g. Whinston 1990; Nalebuff 2004; Bakos and Brynjolfsson 2000), as witnessed by ongoing antitrust challenges to Microsoft's bundling of software applications (e.g. its Internet browser, media player) with its dominant Windows operating system (Mitchener and Kanter 2004).

An influential theoretical literature, however, suggests bundling may arise in many contexts to sort consumers in a manner similar to second-degree price discrimination (Stigler 1968; Adams and Yellen 1976; McAfee et al. 1989). When consumers have heterogeneous tastes for several products, a monopolist may bundle to reduce that heterogeneity, earning greater profit than would be possible with component (unbundled) prices. Recent research suggests similar effects in oligopoly markets (Stole 2003). Bundling—like price discrimination—allows firms to design product lines to extract maximum consumers surplus. While firms clearly benefit in this case, consumer welfare can fall, particularly when bundling requires consumers to purchase products in which they have little interest (Bakos and Brynjolfsson 1999; Armstrong 1996). Textbooks in intermediate microeconomics, industrial organization, and business strategy regularly describe the discriminatory effect of product bundling and suggest its use as an effective business strategy (e.g., Varian 2003; Carlton and Perloff 2001; Saloner et al. 2001).

Despite this variety of motivations to bundle, there is little systematic evidence of its determinants in particular settings. The purpose of this paper is to test the discriminatory incentives to bundle and quantify their importance in the cable television industry. This industry provides a natural environment for several reasons. First, it is considered the canonical example of the discriminatory theory at work (e.g., Wildman and Owen 1985; Chae 1992; Salinger 1995; Bakos and Brynjolfsson 1999; and Armstrong 1999). Second, cable services are fundamentally bundles of various types of television networks and there is considerable heterogeneity in the size and contents of these bundles across cable markets. This provides the variation necessary to test the theory.

Finally, the bundling decisions of cable systems have recently drawn significant media attention as the Federal Communication Commission (FCC) considers mandated unbundling—or "à-la-carte" pricing—as a possible solution to ever-rising cable prices.[1]

We begin by reviewing the literature describing the discriminatory incentives to bundle and discuss the testable implications of the theory. Simple models with two goods demonstrate how bundling reduces consumer heterogeneity and increases profits when marginal costs for bundle components are low. Recent models with more than two goods, relying on statistical laws of large numbers, demonstrate the generality of this conclusion under certain conditions. The primary testable implication of the discriminatory theory is that demand for product bundles should become more elastic as products are added to the bundle. This effect is idiosyncratic to the discriminatory theory and cannot be generated by alternative incentives to bundle. Using a dataset on a cross-section of cable markets in 1996, we estimate the demand for bundles of widely available cable television services and test whether the addition of each of the top-15 cable television networks to a service bundle has the effect predicted by the theory.

The results yield qualified support for the discriminatory theory: adding six of the top fifteen cable television networks to program bundles significantly increases the elasticity of cable demand (and never significantly reduces it). Moreover, this effect is concentrated among specialty program networks, bundle components that are likely to most reduce consumer heterogeneity. To quantify the profit and welfare implications of these results, we specify a simple model of bundle demand with normally distributed tastes and calibrate it to our estimates of cable demand. The results are suggestive of the discriminatory power of bundling. While bundling of general-interest networks is estimated to have no discriminatory effect, the heterogeneity reduction associated with bundling an average top-15 *special-interest* cable network is estimated to increase profits and reduce consumer welfare, with an average effect of 4.7% (4.0%). On balance, total welfare *increases*, with an average effect per special-interest network of 2.0%. These results suggest (but do not guarantee) that consumers might benefit from having such networks available on an à-la-carte basis.

2 The discriminatory incentives to bundle

2.1 The case of two goods

Most of the discriminatory bundling literature has focused on the incentives to bundle two goods. Adams and Yellen (1976) formalize the seminal work of Stigler (1963) and present examples where bundling is more or less profitable

[1] See Reuters (2003), GAO (2003), Squeo and Flint (2004), FCC (2006).

than component (unbundled) sales. Schmalensee (1984) and Salinger (1995) extend the analysis to the case of normal and uniform tastes.

The primary benefit of bundling to come out of this literature is that it reduces heterogeneity in consumer tastes, permitting greater consumer surplus extraction than would be possible with component (unbundled) prices. Heterogeneity reduction, however, is not sufficient. In more general settings, when bundled sales are preferred to component sales depends on three critical features of preferences and costs. First is the extent of heterogeneity reduction possible from bundling. This increases with the negative correlation in preferences for bundle components (Schmalensee 1984).[2] Second is the level of marginal costs for components. Since bundling requires consumers purchase all goods, some below-cost sales of components can result, reducing the gains from bundling. This becomes more likely the higher are marginal costs relative to the mass of consumer preferences. Third is that bundling requires firms charge a single price. When consumer tastes for components differ considerably, bundling is less attractive than component sales as it permits fewer instruments (prices) to capture consumers' surplus.[3]

2.2 More than two goods

Recent papers by Bakos and Brynjolfsson (1999) and Armstrong (1999) extend the analysis of bundling to consider multiple goods. The following model based on Bakos and Brynjolfsson (1999) highlights the incentives in this case and forms the basis for the empirical tests conducted in this paper.

The Bakos and Brynjolfsson model Suppose there are n discrete components supplied by a monopolist and offered to consumers in a single bundle at price p. Let consumers differ in their preferences (willingness-to-pay) for each of these components, given by a type vector, $v_i = (v_{i1}, \ldots, v_{in})$ and let the utility function of each consumer take a simple additive form:[4]

$$u(v_i, p) = \left\{ \begin{array}{ll} \sum_{c=1}^{n} v_{ic} - p & \text{if she buys the bundle,} \\ 0 & \text{else} \end{array} \right\} \quad (1)$$

If the monopolist cannot observe any individual consumer's taste vector, v_i, but knows the distribution of tastes in the population, it may appear to be sub-optimal to offer a single bundle at a single price. The optimal tariff can be

[2] Negative correlation, however, is not necessary for bundling to be profitable (McAfee et al. 1989).

[3] McAfee et al. (1989) extend the analysis of Adams and Yellen (1976) to consider mixed bundling, the offering of *both* component and bundled sales, and show it always yields (weakly) greater profits than pure bundling. The reason for this is clear: it maintains the benefits of bundling (if any) and strictly increases the number of prices available to capture surplus. Despite this fact, mixed bundling is relatively uncommon, perhaps due to the added administrative costs associated with offering both bundled and component goods.

[4] Note that in this specification there are no income effects nor any complementarity or substitutability in demand, two issues that we discuss further below.

quite complex and difficult to calculate, however, even for simple preference structures (Armstrong 1996; Rochet and Chone 1998; Rochet and Stole 2000).

Instead, authors have relied on arguments based on Laws of Large Numbers (LLNs) to characterize firm behavior (Armstrong 1999; Bakos and Brynjolfsson 1999). In particular, note that bundling aggregates (averages) consumer tastes for bundle components. If tastes for components are not too positively correlated, when bundles are large LLNs imply the distribution of preferences for the bundle becomes more concentrated as n increases.

Bakos and Brynjolfsson (1999) (BB) formalize this logic in the following proposition. Let $x_{in} \equiv \frac{1}{n}\sum_{c=1}^{n} v_{ic}$ be the per-good valuation for consumer i of a bundle of n goods and let μ_n be its mean. Let p_n^*, q_n^*, and π_n^* be the profit-maximizing per-good prices, market share, and profits for a bundle of n goods. Assume the following conditions hold:

Assumption A1: The marginal cost for all components is zero.

Assumption A2: Consumer valuations v_{ic}, $c=1,\ldots,n$ are uniformly bounded with continuous density functions, non-negative support, means μ_c and variances σ_c^2, are not perfectly correlated, and are stationary in the wide sense for all n.[5]

Assumption A3: Consumers have free disposal.

Assumption A4: Single-Crossing of Cumulative Distributions Condition (SCDC). The distributions of consumer valuations is such that

$$\text{Prob}[|x_n - \mu_n| < \epsilon] \leq \text{Prob}[|x_{n+1} - \mu_{n+1}| < \epsilon]$$

for all n and ϵ.

BB Proposition 3 Monotonic Bundling Profits. Given Assumptions A1, A2, A3, and A4, if $\pi_{\hat{n}}^* > \pi_1^*$ and $p_{\hat{n}}^* < \mu_n$, then bundling any number of goods $n \geq \hat{n}$ will monotonically increase the seller's profits, compared to selling them separately.

Proof See Bakos and Brynjolfsson (1999, Appendix).[6] □

The essence of the proof follows from the properties of Laws of Large Numbers. As bundle size increases, the distribution of the average WTP in the population of consumers converges to a single point (the limit of the average μ_n). In the limit, all consumers have the same WTP, the monopolist

[5]Bakos and Brynjolfsson (1999, Footnote 21): "A sequence $\{v_i\}$ is called stationary in the wide sense if $E\{|v_i|^2\} < \infty$ for all i, and the covariance $\text{Cov}(v_{s+i}, v_s)$ does not depend on s. This condition is satisfied, for example, if all v_i are identically distributed with finite mean and variance, and $\rho_{i,j} = \rho^{|i-j|}$ for some ρ in (0,1), and for all i and j."

[6]Strictly speaking, this proposition extends Bakos and Brynjolfsson (1999) Proposition 3 to allow for correlation in tastes. As stated, it is correct, as A4 is strong enough to allow for the more general correlation structure.

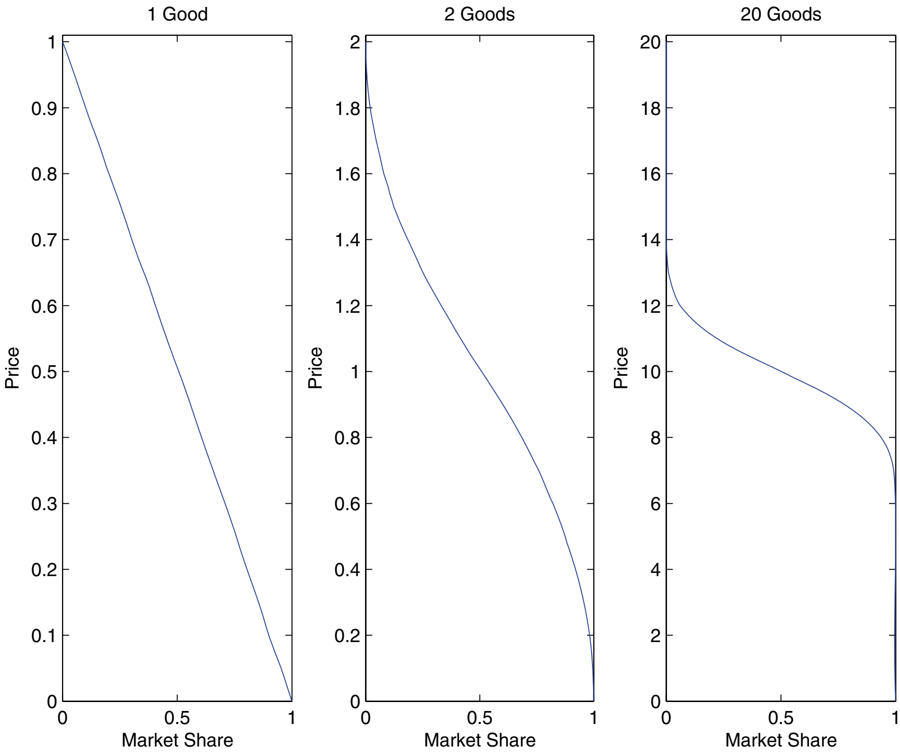

Fig. 1 Demand by bundle size

optimally prices this amount, and earns the maximum possible profit with zero consumer surplus and deadweight loss.

An example of the power of a LLN for the bundle demand curve is demonstrated in Fig. 1, adapted from Bakos and Brynjolfsson (1999). For the case of independent uniformly distributed tastes (i.e. independent linear demand for components), the figure presents the demand for a bundle of size 1, 2, and 20. As bundle size increases, there are fewer extreme tastes, corresponding to an increasingly flat demand curve and greater consumer surplus extraction.[7] Similar effects obtain for other distributions.

Each of the stated assumptions is important to this result, but likely to hold for the cable television services we examine in this paper. Assumption A1 is strong: with zero marginal costs (and A3, free disposal), adding components is always profitable (and efficiency enhancing). While marginal costs are not strictly zero in cable television markets, they are likely low enough that systems rarely end up bundling products which a significant mass of households value

[7] This may seem counter-intuitive. For a fixed level of demand (e.g. rotate a linear demand curve around its intersection with the quantity axis), a monopolists profit is higher the more inelastic is demand. Bundling, however, simultaneously *shifts out* and flattens the aggregate demand curve.

less than their cost.[8] Assumption A2 is weak by comparison: the requirement that preferences are not "stationary in a wide sense" only rules out strong forms of positive correlation in tastes for components.[9] Assumption A4 simply ensures that the homogenizing effect of a LLN applies in small samples, i.e. that the distribution of preferences for bundles becomes more concentrated *for each incremental component*.[10]

2.3 Testing the discriminatory theory

We test two primary implications of the discriminatory theory in this paper: Does the bundle demand curve become more elastic with increases in bundle size? And does it increase more for components that are most negatively correlated with components of existing bundles?

We formalize these tests in the following propositions. Let $\tilde{p}_n^*$ be the profit-maximizing total price for a bundle of size n (i.e. $\tilde{p}_n^* = np_n^*$).[11]

Proposition 1 *Let* $\tilde{\epsilon}^n(\tilde{p}_n^*) \equiv -\frac{\partial q^n(\tilde{p}_n^*)}{\partial p}\frac{\tilde{p}_n^*}{q_n(\tilde{p}_n^*)}$ *be the (absolute value of the) elasticity of demand for a bundle of size n evaluated at its profit-maximizing price. Under symmetry and the assumptions of BB Proposition* 3 *above,* $\tilde{\epsilon}^{n+1}(\tilde{p}_n^*) > \tilde{\epsilon}^n(\tilde{p}_n^*)$*, i.e. the elasticity of demand for a bundle of size* $n+1$ *evaluated at the profit-maximizing price for the size-n bundle is more elastic than the elasticity of demand for a bundle of size n evaluated at that price.*[12]

---

[8]Recent exceptions may include sport networks. We discuss this issue in more detail in the next subsection.

[9]Both Armstrong (1999) and Bakos and Brynjolfsson (1999) note that this assumption might be violated if preferences for components are correlated with one or more underlying variables (esp. income). For example, high-income households might be willing to pay more than low-income households for all cable networks. In this case, Armstrong (1996) shows that the "almost-optimal" tariff would be a *menu* of cost-based two-part tariffs (with the menu designed to sort households by income) and Bakos and Brynjolfsson (1999) note that the properties of their propositions can be restored by *conditioning* on income. We include demographic variables (including income) as demand shifters for each offered cable bundle in our econometric model in part to account for this concern. See Crawford and Shum (2006) and Xiao et al. (2006) for analyses that allow for menus of product bundles.

[10]As noted in BB, if preferences for components are i.i.d., the single-crossing condition (SCDC) is not particularly restrictive. It is satisfied for most common demand functions, including the logit specification used in our empirical tests. If, however, preferences for components are positively correlated or differ dramatically in their dispersion (as measured by σ), the assumption may not hold and there is no guarantee bundling reduces heterogeneity and increases profits. This provides power for our empirical tests.

[11]For convenience, Bakos and Brynjolfsson (1999) focus on the across-component average willingness-to-pay for a bundle. The aggregate effects of bundling, however, apply equally to demand for the whole bundle: under A1–A3, mean WTP for the bundle increases at rate n while the standard deviation increases at rate $\sqrt{n}$. The coefficient of variation for the bundle, $\sigma_{\text{bun}}/\mu_{\text{bun}}$, therefore *decreases* at rate $\sqrt{n}$. We consider both the per-good and aggregate implications in Appendix 1.

[12]Alternative formulations of this proposition are possible (e.g. evaluating the elasticity at a given quantity, at the mean of consumer preferences, or evaluating each elasticity at its profit-maximizing price). We chose the given form as that provides the most natural empirical implementation. See footnote 25 for the implication of this choice for our empirical tests.

Proof See Appendix. □

The essence of the proof follows the intuition in Fig. 1. As bundles become larger, preferences become more concentrated. As such, incremental changes in price have larger quantity effects and the elasticity increases.

Proposition 2 *Let $\tilde{\epsilon}^n(\tilde{p}_n^*)$ be defined as in Proposition* 1 *above and let $\rho_{n\tilde{-}1,n}$ measure the correlation in preferences between a bundle of $n-1$ products and the nth product, i.e. $\rho_{n\tilde{-}1,n} \equiv corr(x_{i,n-1}, v_{in})$. Then $\frac{\partial \tilde{\epsilon}^n(\tilde{p}_n^*)}{\partial \rho_{n\tilde{-}1,n}} < 0$, i.e. the demand for a bundle of size n is more elastic the more negative is the correlation between the new product and the existing bundle.*

Proof See Appendix. □

The essence of the proof follows the intuition of the impact of correlation on the variance of a sum of random variables. If $V(A+B) = VA + VB + 2COV(A, B)$ for two random variables A and B, the more negatively correlated are A and B, the lower is $V(A+B)$. When bundling, lower variance means greater heterogeneity reduction which means more concentrated tastes which means more elastic demand.

Discussion: Costs and Concept Assumption A1 underlying Bakos and Brynjolfsson (1999) Proposition 3 (and therefore our tests) requires marginal costs for components to be zero. Recall also that marginal costs are important, as the higher are marginal costs for components the less profitable and efficient is bundling as it can require sales of components to consumers who value them less than their cost. As television programming is broadcast by satellite across the entire country, the marginal cost to program networks of distributing television signals is effectively zero. Networks do, however, charge affiliate fees (in units of dollars per subscriber per month) to cable systems for the right to carry that network on a given cable system. For basic cable networks (e.g. MTV, ESPN), these vary from effectively nothing (with the majority of network revenues coming from advertising) to a 1993 high of $0.60 per subscriber per month for ESPN (Kagan World Media 1998). For (unbundled) Premium Cable Networks (e.g. HBO), they can be as high as several dollars per subscriber per month. In part for this reason, we will examine the demand for bundles of basic cable networks in this paper. Hazlett and Spitzer (1997, Tables 5–10) estimate 1993 average spending per (basic) subscriber of $3.21 per month.[13] While not strictly zero, this is likely to be sufficiently below the mass of support of the distribution of willingness-to-pay for cable bundles to make the theory applicable.[14]

[13] Versus an average 1995 price of $17.33 for the most popular bundle in our data.

[14] Indeed if it were not so, one might expect cable systems to optimally unbundle relatively expensive basic networks as they do premium networks. Footnote 20 describes a desire for movement in this direction for sports networks.

Note also that the increase in the cable demand elasticity is an idiosyncratic prediction of the discriminatory theory and cannot be generated by alternative incentives to bundle. The leading alternative is complementarity in cost, or economies of scope. As described in the next section, cable distribution technology suggests the presence of important cost complementarities. So too may administrative and marketing costs exhibit complementarities. Such complementarities should not, however, impact cable demand. Similarly, complementarity in demand is unlikely to yield effects like that predicted by the discriminatory theory and the extension of market power is unlikely to be important when all cable/satellite firms have nondiscriminatory access to bundle components.[15] The implications described above are therefore unique to the discriminatory theory.

That being said, finding support for the discriminatory theory does not necessarily disprove other incentives to bundle in cable television markets. Indeed, suppose that the primary reason to bundle in cable was complementarity in cost. If so, and the assumptions underlying the Proposition above also happened to hold, there would be both cost and demand (heterogeneity reduction) benefits of bundling. Our purpose here is therefore not to ask if discriminatory incentives are *the* reason cable systems bundle; rather we wish to ask if discriminatory incentives could be *one* reason cable systems bundle and, if so, to try to quantify their importance to profit and welfare. We describe the specifics of our test after introducing patterns of bundling in the cable industry.

3 Bundling in the cable television industry

The cable television industry is considered a canonical example of discriminatory bundling in the economics literature (Wildman and Owen 1985; Salinger 1995; Bakos and Brynjolfsson 1999). In this section, we describe patterns of bundling in the industry and characterize the institutional and regulatory constraints placed on system's bundling decisions.

Cable services: bundles of program networks Cable television systems choose a portfolio of television networks, bundle them into services, and offer these services to consumers in local, geographically separate, cable markets. All cable systems offer four main types of program networks. *Broadcast networks* 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 networks—ABC, CBS, NBC, and FOX—as well as public and independent television

[15]The impact of complementarity in demand on the cable elasticity would depend on how complementarities varied within the distribution of households. If (as seems likely) preferences in high-WTP households exhibited stronger complementarities than in low-WTP households, the bundle demand curve would likely shift out and rotate clockwise with additions to the bundle, decreasing the (absolute value of the) elasticity of cable demand.

stations. *Cable programming networks* are advertising- and fee-supported general and special-interest networks distributed nationally to systems via satellite. Examples include some of the most recognizable networks associated with cable, including MTV, CNN, and ESPN. *Premium programming networks* are advertising-free entertainment networks, typically offering full-length feature films. Examples include equally familiar networks like HBO and Showtime. *Pay-per-view networks* are specialty channels devoted to on-demand viewing of high-value programming, typically offering the most recent theatrical releases and specialty sporting events.

Systems exhibit moderate differences in how they bundle networks into services. Broadcast and cable programming networks are typically bundled and offered as *basic service* while premium programming networks are typically unbundled and sold as *premium services*.[16] In the last decade, systems have begun to further divide basic service, offering some portion of their cable networks on multiple services, called *expanded basic services*.[17] For either basic or expanded basic services, consumers are not permitted to buy access to the individual networks offered in bundles; they must instead purchase the entire bundle.

Institutional and regulatory constraints on bundling in cable What drives patterns of bundling in the cable industry? While the focus of this paper are the discriminatory incentives to bundle, we briefly describe several institutional and regulatory constraints that impact firms' bundling decisions.

The first is complementarity in cost, or economies of scope. The least cost method of providing any cable service is to bundle all the programming. This is due to the underlying technology of video program distribution: all television networks are transmitted to each customer's home. It is *unbundling* networks that is costly, requiring methods to prevent consumption by non-subscribers. The technology to do this has changed over time, implying the production technology of a given cable system can significantly impact the cost (or even the feasibility) of providing alternative bundles of programming.[18] Why then do systems unbundle at all? The discriminatory theory suggests an answer. Systems may well trade off the benefits of heterogeneity reduction from bundling against the cost of sales of components to households that value them less than their cost. Depending on their technology, systems may

[16]Premium networks have recently begun "multiplexing" their programming, i.e. offering multiple channels under a single network/brand (e.g. HBO, HBO 2, HBO Family, etc.).

[17]With the rise of digital cable, many cable systems now offer more service tiers. These were not available, however, in the time period we study.

[18]Early methods to block consumption relied on electromechanical "traps" placed at the link between the household and the cable distribution system. Most (but not all) systems now offer "addressable" converters which control access via electronic communication with the cable headend.

optimally unbundle high-cost (e.g. premium) networks but not low-cost (e.g. cable) networks.[19] This has been the historical pattern in the industry.

Why then do not systems also unbundle basic and expanded basic services? The feasibility of low-cost unbundling accompanying addressable converters suggests this is now feasible. A recent report by GAO (2003) suggests some answers. The first is that not all consumer opt for addressable converters, even when offered by their system. Uniform deployment of converters, while likely in the long-run, could be costly at present. This raises the costs of unbundling. The second is that *networks* do not want to be unbundled.[20] The average cable network earns about 50% of its revenue from advertising (GAO 2003). Unbundling would clearly reduce the set of consumers that could watch a network and likely reduce the number that do watch. This would plausibly reduce advertising revenues and require uncertain increases in license fees to compensate.[21]

A second influence on systems bundling decisions is regulation. While the specific content of any cable service may not be regulated on First Amendment grounds, the 1992 Cable Act introduced several rules that impact patterns of bundling in the industry. The first required the creation of a basic tier of service containing all offered broadcast and public-interest programming carried by the system, as well as cable programming networks (at the discretion of the system). In addition, price regulations introduced by the 1992 Act set different rules for bundled versus unbundled (a-la-carte) services, meaning systems had an incentive to change their mix of offered services and the programming provided on them (Hazlett and Spitzer 1997; Crawford 2000). While this affected the number of services provided by some systems, changes in the networks provided were relatively minor (Crawford 2000). This is perhaps not surprising, as *tiering* decisions (e.g. how many, if any, expanded basic Services are offered) are often made by a cable system's corporate parent, or multiple system operator (MSO), while *carriage* decisions (i.e. what networks to offer on those services) are typically made by the local system.

In summary, many institutional and regulatory factors encourage bundling in cable markets, particularly for basic and expanded basic services. Most of these, however, impact the costs associated with unbundling networks are therefore complementary to reductions in consumer heterogeneity implied by the discriminatory theory. In the next section, we introduce the data and econometric model used to test the implications of the theory.

---

[19] Cable networks typically charge moderate fees to cable systems, ranging from nothing to $1.00 or more per subscriber per month. By contrast, premium networks charge higher fees, up to several dollars per subscriber per month for HBO.

[20] Witness the recent disputes between ESPN and Cox and the NFL Network and Time Warner/Cablevision over the possible unbundling of sports networks (e.g. www.keepespn.com, www.makethemplayfair.com, Gentile 2004 and Thompson 2006).

[21] For example, Direct Broadcast Satellite providers of multi-channel programming in competition with cable systems face no technological constraint but also engage in widespread bundling.

4 Empirical specification

4.1 Data

I've compiled a market-level dataset on a cross-section of United States cable systems to test the discriminatory theory. The primary source of data is Warren Publishing's Television and Cable Factbook Directory of Cable Systems. The data for this paper contains all cable systems recorded in the 1996 edition of the Factbook for which complete information was available. This yielded 1,159 systems.[22]

Table 1 presents some summary statistics for these systems. While all systems offer a basic service, Table 1 shows that slightly more than a third of systems offer expanded basic services. Of these, most offer just one Expanded Service. Aggregating over all basic and expanded basic services, systems typically offer almost 6 broadcast networks and more than 17 cable networks. When modeling the demand for cable services, it is important to identify the specific networks offered to households (Crawford 2000). We do so according to the size of their potential audience: the top 15 cable programming networks available in the United States in 1998 are listed in Table 2. Table 3 reports summary statistics regarding the allocation of these networks across services. The first column reports the proportion of systems in the sample that carry each of the top-15 cable programming networks on *any* basic service. The remaining columns examine the proportion of systems that carry each top-15 cable network on each basic or expanded basic service. Several interesting patterns emerge. First, the majority of the top 15 networks are offered on some service by the majority of systems. Some of the most popular networks, for example WTBS, CNN, and ESPN, are available on over 95% of systems. Systems differ, however, in how they allocate these networks among basic and expanded basic services. While some, like CSPAN and QVC, are almost exclusively offered on basic, others, like TNT and TNN, are often found on expanded services. Importantly, there is significant heterogeneity both in the carriage of networks across systems, as well as in their allocation to basic and expanded basic services.[23]

[22] While there are over 11,000 systems in the sample, persistence in non-response over time as well as incomplete reporting of critical variables required imposing a large number of conditions in order for a system to be included in each sample. Missing information on prices, quantities, and reporting dates were responsible for the majority of the exclusions.

[23] At the time of the sample data, the heterogeneity in carriage and allocation of a given network to a given service across systems was largely driven by two factors: variation in the tiering strategies taken by different MSOs (with a given strategy typically implemented in most but not all of an MSO's systems) and variation in local tastes for programming (influencing especially carriage decisions). See footnote 32 for more on this issue.

Table 1 Summary statistics

Variable	Mean	SDev	Min	Max
Expanded basic services				
Any exp. basic svcs.	0.37	0.48	0.00	1.00
One exp. basic svc.	0.23	0.42	0.00	1.00
Two exp. basic svcs.	0.14	0.35	0.00	1.00
Market shares				
w_{Basic}	0.68	0.15	0.18	0.99
$w_{\text{Exp. I}}$	0.23	0.31	0.00	0.97
$w_{\text{Exp. II}}$	0.07	0.20	0.00	0.93
Programming				
Broadcast networks				
Over-the-air	2.60	1.42	0.00	8.00
On cable	5.84	2.03	0.00	13.00
Cable networks				
Top-15, any basic	10.46	3.29	1.00	15.00
Other than top-15, any basic	6.65	5.09	0.00	32.00
Total, any basic	17.11	7.84	1.00	47.00
Other basic channels	13.92	10.47	0.00	56.00
Total channels, any basic	36.86	13.49	5.00	107.00
Prices				
p_{Basic}	17.33	4.82	4.42	37.07
$p_{\text{Exp. I}}$	2.87	5.09	0.00	24.08
$p_{\text{Exp. II}}$	0.63	1.64	0.00	14.73
Total price, all basic	20.83	4.40	4.97	59.72
Instruments				
Homes passed (000s)	5,076	17,443	48	275,394
MSO subscribers (0,000s)	79.28	194.99	0.00	1200.00
Affiliation	0.09	0.29	0.00	1.00
Channel capacity	39.22	13.93	6.00	110.00

Sample is 1,159 US cable systems. Cable data from *The Cable and Television Factbook*, vol. 64 (1996) by Warren Publishing. Over-the-air broadcast programming defined as "Significantly Viewed" broadcast stations by county from *Cable and Station Coverage Atlas*, 1987 by Television Digest. All systems offer Basic Service and up to two expanded basic services, indexed by I and II. Market shares defined as subscribers divided by homes passed, defined as households able to purchase cable services from each system. Top-15 Networks defined in Table 2. Multiple System Operator (MSO) Subscribers defined as the total subscribers to all systems owned by same firm. Affiliation equals 1 if system owned by MSO with ownership interests in programming networks. Homes Passed and MSO Subscribers measured in thousands, tens of thousands, respectively. See Crawford (2000) for more detail on data sources.

4.2 An econometric model of demand for cable television services

Recall a cable service is a bundle of television networks. The goal of the demand model is to estimate the impact of changes in bundle size on the elasticity of demand for these services. To do this, I specify a logit model of demand for basic and expanded basic cable services across markets. Let the aggregate demand for each of the basic and expanded basic services, s, offered by a cable system in market n be given by

$$\log\left(\frac{w_{sn}}{w_{0n}}\right) = X'_{sn}\beta + D'_n\theta_s + \left(\alpha_s + X'_{sn}\gamma + D'_n\theta^{\text{p}}\right) p_{sn} + \nu_{sn} \qquad (2)$$

Table 2 Top-15 cable programming networks

Rank	Network	Subscribers (millions)	Programming Format
1	TBS Superstation	77.0	General interest
2	Discovery Channel	76.4	Nature
3	ESPN	76.2	Sports
4	USA Network	75.8	General interest
5	C-SPAN	75.7	Public affairs
6	TNT	75.6	General interest
7	FOX Family Channel	74.0	General interest/kids
8	TNN (The Nashville Network)	74.0	General interest/country
9	Lifetime Television	73.4	Women's
10	CNN (Cable News Network)	73.0	News
11	A&E	73.0	General interest
12	The Weather Channel	72.0	Weather
13	QVC	70.1	Home shopping
14	The Learning Channel (TLC)	70.0	Science
15	MTV: Music Television	69.4	Music

Data on network subscribers from NCTA (1998). Data on programming formats from individual network promotional material (available from http://www.ncta.com), NCTA (1998), or industry sources.

where s indexes {basic service, expanded basic service 1 (if offered), expanded basic service 2 (if offered)}, 0 indexes the purchase of no cable service, w_{sn} is the market share of service s in market n, X_{sn} indexes the programming provided on service s in market n, D_n measures demographic attributes of the population in market n, p_{sn} measures the price of service s in market n, and ν_{sn} measures unobserved shocks to demand for service s in market n. α_s, β, and γ measure, respectively baseline aggregate price sensitivity for cable service s, the impact to demand from the carriage of the program networks in X_{sn}, and the impact to aggregate price sensitivity from the carriage of the networks in X_{sn}. θ_s and θ^{p} measure differences in the tastes for cable services arising from differences in demographic characteristics of cable markets. The market share for service s in market n, w_{sn}, is defined as the number of subscribers to that service divided by the number of homes passed by the cable system, where homes passed are the number of households accessible by a cable system's distribution network.[24] Demand shifters are denoted by D_n and include the Designated Market Area (DMA) rank and its square, measuring the strength of the local television market, median income and its square, the percentage of the population aged 5 to 18, and the percentage of the population with any college experience. Also included are region dummies to control for taste differences across regions and Expanded Service dummies in the basic and expanded service equations.

[24]This is a reliable measure of market size as it defines the set of households available to purchase cable services from each system.

Table 3 Summary statistics, top-15 cable networks

Services	Any basic	Basic	Expanded basic I	Expanded basic II
TBS	0.98	0.77	0.10	0.11
Discovery	0.83	0.55	0.24	0.05
ESPN	0.98	0.79	0.19	0.00
USA Network	0.87	0.60	0.25	0.01
C-SPAN	0.42	0.36	0.06	0.00
Top-5	4.09	3.07	0.84	0.17
TNT	0.81	0.55	0.20	0.06
Family	0.92	0.69	0.19	0.04
TNN	0.93	0.63	0.25	0.05
Lifetime	0.52	0.36	0.15	0.00
CNN	0.96	0.67	0.25	0.04
Top-10	8.22	5.97	1.88	0.36
A&E	0.52	0.39	0.13	0.00
Weather	0.47	0.30	0.15	0.02
QVC	0.50	0.47	0.04	0.00
TLC	0.24	0.19	0.05	0.00
MTV	0.51	0.38	0.13	0.00
Top-15	10.46	7.70	2.37	0.39
Other Cable Nets.	6.65	4.85	1.53	0.28
Total Cable Nets.	17.11	12.54	3.90	0.66

Reported are the proportion of sample systems carrying each top-15 network on basic service, expanded basic service I, or expanded basic service II and corresponding average number of networks offered. First column reports carriage on *any* offered service (any basic). Remaining columns disaggregate carriage by service.

From above, ν_{sn} measures unobserved shocks to demand for service s in market n. While there are many plausible reasons for demand unobservables, the two most likely in cable television markets are unobserved (to the econometrician) attributes (quality) of service s in market n and unobserved tastes for service s in market n not captured by X_{sn} and D_n in the empirical specification. We discuss the implications of these unobservables after introducing the instruments in the next subsection.

Econometric tests of the discriminatory theory Propositions 1 and 2 in Section 2 described the comparative statics results we will test in this section. Proposition 1 stated that the bundle demand curve becomes more elastic with increases to the size of the bundle. Proposition 2 stated that this increased elasticity should be greater the more negatively correlated is the new product with existing products in the bundle.

To test the discriminatory theory, we construct hypothesis tests analogous to each of these implications. The key explanatory variables in these hypothesis tests are those contained in the vector of programming networks, X_{sn}, bundled on service s. Under Proposition 1, if bundling facilitates price discrimination, adding a network in X_s to a service bundle should make demand more elastic. Under Proposition 2, this effect will be greater the greater is the negative correlation between that network and other networks in the bundle.

Ideally, we would like to measure the impact on the demand elasticity of adding *each* network offered on any cable system. There are far too many networks, however, for this to be practical. Instead, we measure the impact on the demand elasticity of adding each of the top 15 cable networks, aggregating the remaining cable networks into a single category, "other cable networks." This implies 15 separate tests of the discriminatory incentives to bundle, one corresponding to each top-15 network.

The essence of the empirical strategy is described by Fig. 2. Represented there are two cable demand curves: one for a bundle of networks excluding MTV and one with a bundle of networks including MTV. Suppose the decision to offer MTV is exogenous (more on this later). We wish to estimated the difference in the level (β) and slope (γ) of the two demand curves, summarizing their effects by comparing the estimated elasticity of each, ideally evaluated at $p^*_{\text{No MTV}}$.[25] We then repeat this for the other top-15 networks in the bundle.

Formally, let $\tilde{\epsilon}^{Bn}(X_{Bn}, p_{Bn}, D_n)$ be the elasticity of demand for basic cable service in market n, evaluated at the vector of observed networks and price offered in market n. let x_{cBn} be one of the components of X_{Bn}, $c = 1, \ldots, 15$. Let X^{c0}_{Bn} be equal to X_{Bn} except that we have set the c-th element of that vector to 0, i.e. $x_{cBn} = 0$. Let X^{c1}_{Bn} be equal to X_{Bn} except that we have set $x_{cBn} = 1$. The tests of Proposition 1, that adding a network to a service bundle makes it more elastic, are then given by $H^{P1}_0 : \tilde{\epsilon}^{Bn}(X^{c0}_{Bn}, p_{Bn}, D_n) > \tilde{\epsilon}^{Bn}(X^{c1}_{Bn}, p_{Bn}, D_n)$, for $c = 1, \ldots, 15$. If we can reject H^{P1}_0 for any c, we can conclude that adding network c makes the bundle demand curve more elastic. The tests of Proposition 2 are analogous: that the difference in elasticity in H^{P1}_0 is larger for networks that correlate negatively with other networks in the bundle.

Which networks might correlate negatively with other bundle elements? Programming formats provide a possible answer. Recall the list of top-15 networks presented in Table 2 categorized networks into formats characterizing the types of programming provided. The broadest distinction in the table is between "general-interest" programming appealing to a wide range of tastes and "special-interest" programming appealing to a narrow range of tastes.[26] This categorization, we claim, is plausibly related to taste (co-)variation. In particular, if (1) different special-interest networks are targeted to different segments of the viewing population (as for example MTV targets young

[25]Evaluating the elasticity at $p^*_{\text{No MTV}}$ turns out to be difficult as we do not observe this price in the data and simulating it would require modeling firms' optimal pricing decisions. Instead, we evaluate the elasticity at the *observed* bundle prices, themselves a weighted average of the optimal No-MTV and MTV bundle prices. If anything, this biases us against finding evidence of the discriminatory theory as that price is too high for the No-MTV bundle (over-estimating its elasticity) and too low for the With-MTV bundle (under-estimating its elasticity). Alternatives (e.g. evaluating elasticities at a given quantity) yielded stronger support for the discriminatory theory but made it more confusing to measure the profit and welfare effects (which necessarily had to be evaluated at profit-maximizing prices).

[26]Examples of the former include TBS, USA, and TNT. Examples of the latter (with their primary type of programming) include ESPN (Sports), CSPAN (Public Affairs), and QVC (Home Shopping).

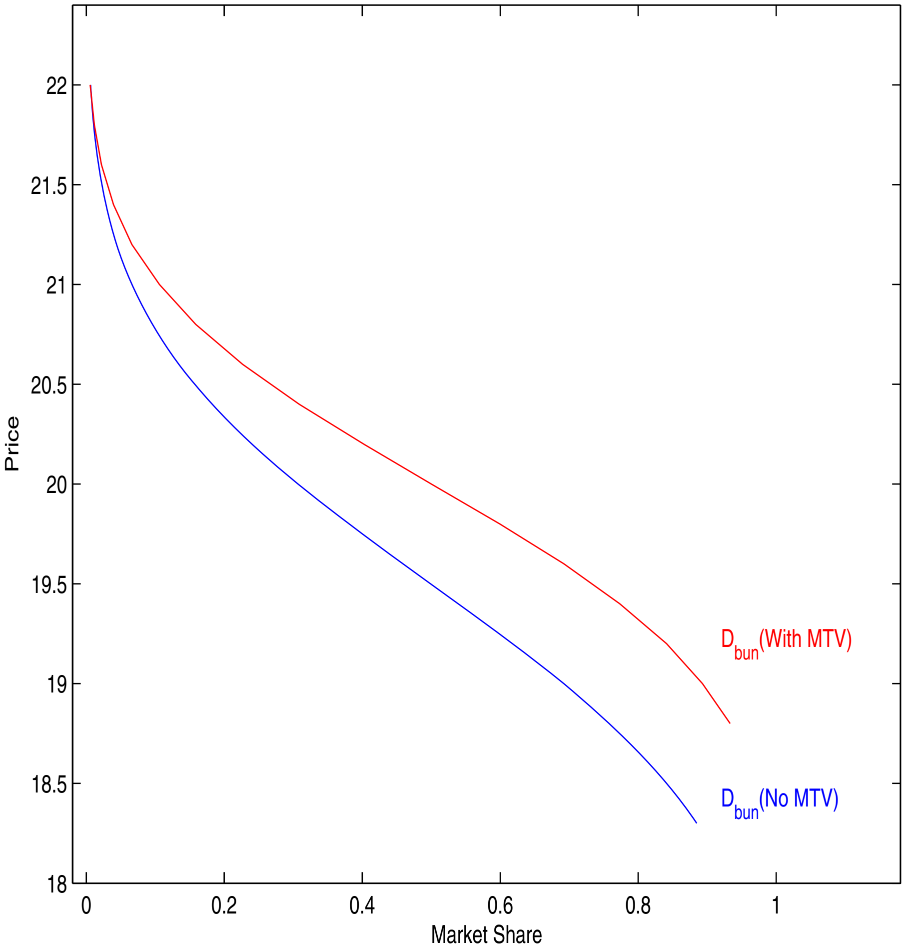

Fig. 2 Bundle demand with and without MTV

adults and Lifetime targets adult women) and (2) tastes of these population segments negatively covary (as for example preferences of young adults and their mothers), one would expect tastes for special-interest networks to be more likely to negatively covary with other networks in an existing bundle than would general-interest networks.[27] If so, the bundling of special-interest networks should increase the cable demand elasticity more than the bundling of general-interest networks. This is our test of Proposition 2.

[27]This is not perfect, however. As noted by a referee, tastes for some pairs of special-interest networks may be independent or even positively correlated (e.g. sports and action networks). We do not dispute that. The claim here is that, *on average*, tastes for a special-interest network and an existing bundle are more likely to be negatively correlated than tastes for a general-interest network and a bundle.

Discussion There are a number of restrictions embodied in Eq. 2. First is the logit assumption. This is made for convenience; in principle, any aggregate demand specification would suffice. Indeed, nested logit and probit models yielded qualitatively similar results.[28] Second is the restriction that some parameters (β, γ) are the same across the services offered in a given cable market (if there are more than one). While in principle there might be different effects from, e.g., adding ESPN to a basic versus expanded basic package, the data were not (and are unlikely to be) rich enough to separately identify effects for networks on each different service tier. The results presented here were robust to specifications that relaxed this assumption.[29]

4.2.1 Instrumental variables

Price instruments Recent developments in the industrial organization literature have emphasized the importance of allowing for endogenous prices when estimating the demand for differentiated products (Berry 1994; Berry et al. 1995). In particular, it is likely that cable systems select prices for their services in market n, p_{sn}, based on information about their products or consumers tastes captured by ν_{sn} that are unobservable to the econometrician. Results from estimation with two sets of price instruments are therefore reported.

The primary marginal costs to a system are the per-subscriber fees paid to programming networks. These costs are considered sensitive competitive information, however, and are not widely available (GAO 2003). The first set of instruments therefore proxies for differences across systems in these costs. The first three cost shifters, homes passed and the number and square of subscribers served by the systems corporate parent (MSO), proxy for system size at the local and national level. They capture differences in the marginal programming cost arising from heterogeneity in bargaining power in the programming market (Noam 1985; Chipty 1995) and are likely to be unrelated to demand shocks.[30] We also include a dummy variable if a system's MSO

[28]Estimation using logit and nested logit models simplifies the estimating equations when some products are bundles that include the contents of other products (Crawford 2000). This and a lack of precision on the nesting parameter(s) in the nested logit model led us to prefer the logit specification.

[29]In particular, estimating using only one bundle per market—whether the lowest or highest quality—yielded similar (though less precise) results.

[30]The logic of this argument is that households have preferences for the set of services offered by cable systems, but are unlikely to care about the size or identity of provider of those services. There is little branding of cable *systems* (in contrast to cable *networks*). This may not be true in the long run, however, if certain providers provide (unobservably) higher quality cable services. The nature of the product in cable helps alleviate this concern. It is easy to enumerate the things that matter to consumers about cable service; most important by far is the set of program networks they can watch and these are observed by the econometrician. Other features (signal quality, customer service, etc.) are likely of second-order importance to demand and regardless unlikely to be correlated with programming cost shifters.

has vertical ties to programming networks. Both Chipty (2001) and Waterman and Weiss (1996) find that systems tend to favor affiliated networks, at least in part because they can purchase programming from their affiliates at its true, very low, marginal cost. A final cost shifter, channel capacity, proxies for the ability for systems to earn reduced rates on bundles of programming networks provided by the same supplier.

We also report specifications based on a second set of instruments that rely more heavily on institutional features of cable program supply. In cable, a system's corporate parent, or MSO, negotiates programming fees on behalf of all its member systems. Individual systems then select the networks to offer given these input prices. As a result, marginal costs across systems within an MSO share common components, a fact that may be exploited in constructing instrumental variables. Because of these common components, differences in prices across systems within an MSO reflect either differences in demand for cable service across markets or idiosyncratic components of cost. If demand shocks for systems owned by a given MSO are not correlated, prices for systems within an MSO outside of a market, n, will be good instruments for prices within market n.[31] We call these instruments 'MSO Prices'.

The use of prices in other geographic markets as instrumental variables has recently been successfully implemented in the market for ready-to-eat cereals by both Hausman et al. (1994) and Nevo (2001). The primary concern in its use is that the assumption of independent demand shocks across markets may not hold, introducing inconsistency into the econometric estimation. Most problematic in cable is geographic concentration of ownership by MSOs. The sorting of people with similar preferences into communities (or regions) introduces the possibility of regional correlation in tastes. We include region dummies in the demand estimation as a first line of defense against such concerns. In addition, We also construct variations of this class of instrument based on prices for systems within an MSO which are located in another state of the country. This necessarily reduces sample size for these specifications, but mitigates concerns that regional demand shocks might bias the results.

Network instruments In addition to setting prices, systems also plausibly select (1) the networks to offer on basic and expanded basic services and

[31]The identification assumption may be described by assumptions on the following generic reduced form for cable prices:

$$p_{sn} = c_{sn} + \epsilon_{sn}$$

where p_{sn} measures the price of good s in market n, c_{sn} is its marginal cost, and ϵ_{sn} measures unmodeled cost and demand shocks to cable prices (including markups). Then if $E(c_{sn}c_{sn'}) \neq 0$ and $E(\epsilon_{sn}\epsilon_{sn'}) = 0$, prices in other markets will be valid instruments for prices in market n. The nature of MSO bargaining for programming networks justifies the correlation in marginal costs across markets within an MSO. The validity of the assumption on ϵ is discussed in what follows.

(2) their allocation across services on the basis of unobserved tastes in each market. If so, then the contents of bundles will be endogenous.[32] This is an important concern often overlooked in the IO literature, where product characteristics are routinely assumed to be exogenous. To address this issue requires one of two approaches: explicitly model systems' bundling decisions or instrument for the content of network bundles. The first approach is computationally intensive and beyond the scope of this paper.[33] Instead, we report results from estimation with instruments for networks as well as prices.

The number of networks offered by systems complicates matters. What is needed are variables that shift the probability that a system carries a given network which are uncorrelated with demand for a cable service including that network. As for prices, cost shifters are best, but finding variables that shift marginal costs for individual networks is a difficult challenge.

The solution implemented here constructs instruments for network carriage based on carriage decisions for other systems owned by the same MSO. Specifically, for each network offered on each service for each system in the sample, we estimate the average likelihood of offering that network on the same service at all other systems owned by the same MSO. This is the same instrumenting strategy described above for the second set of price instruments. As there, it will be a valid instrument as long as demand shocks for systems within an MSO are not correlated. As above, we assess the robustness of this assumption by also constructing a variation of this instrumented based on network carriage for systems within an MSO which are located in another state.

Summary We present results from reduced form regressions of both price and network carriage (RHS endogenous variables) on all included exogenous variables and instruments in Appendix 2. With few exceptions, we can reject the null hypothesis of joint insignificance for the instruments in each reduced form regression. This suggests (but does not guarantee) the instruments have power.

---

[32] This may not be as severe a problem as would appear at first glance. In cable, while local systems typically select what networks to offer, the decision to offer one or more expanded basic services and the decision of where to place networks, *if offered*, among these services is typically made by the MSO (GAO 2003). In addition, as most systems offer most networks (cf. Table 3), the econometric identification of tastes for networks is driven as much by the service on which a network is offered as whether it is carried at all. As such, it is the allocation of networks across services that is the important source of possible correlation with the econometric error, at least for the most popular networks. Since this decision is made by the MSO, it is unlikely to be correlated with tastes for cable in any particular market. This is less true for less popular networks, where endogeneity may therefore be a greater concern.

[33] See Crawford and Shum (2006) for some related work on this problem.

5 Results

5.1 Parameter estimates

Table 4 presents the results from estimation of the model of cable demand. The results are organized in groups of three columns. For each group, the first column presents estimates of β, the impact to the level of demand from adding the reported cable network to a cable service bundle. The second column presents estimates of γ, the impact to aggregate price sensitivity from adding the reported network to a cable service bundle. The third column combines these effects and presents the mean impact of adding the reported network to the basic demand elasticity. This is calculated in each market by calculating the basic cable elasticity with and without the network in question, evaluated at the observed values of other offered networks. The average across markets for each network is reported. Bootstrap standard errors are in parentheses.

For each specification, results are reported for each of the top-15 cable networks grouped into general-interest and special-interest networks according to the classification in Table 2. Also reported are an average effect for each group. for all the top-15 networks, and the estimated price sensitivity, $\hat{\alpha}_s$, for each of basic service, Expanded Basic I and Expanded Basic II. Not reported but present in the estimation are parameter estimates for the constant, Expanded Service dummies in the basic and expanded equations, broadcast programming, other satellite networks (in levels), bundle size (in slope), and demographic variables.

Baseline results The first group of columns, labelled (1) OLS, presents OLS estimates of the demand system, while the second group, labelled (2) IV: MSO Prices, presents Instrumental Variable estimates using MSO prices as instruments. Examining the bottom of the table, instrumenting for prices increases the estimated price sensitivity for cable: each of the estimated price coefficients moves in the expected direction and estimated elasticities (not reported) increase by approximately 50%. For the remainder of this section, we focus on the Instrumental Variable results from the fourth-sixth columns.

While we report both the level (β) and slope (γ) effects, these do not have the typical "structural" interpretation common in differentiated product demand estimation (e.g. Berry et al. 1995; Nevo 2001; Petrin 2003). Instead, for each top-15 cable network, they (jointly) capture the impact to the cable bundle demand curve of the inclusion of that network in the bundle. The tests of the discriminatory theory aggregate these effects and focus on the impact of the carriage of each network on the (basic) cable demand elasticity.

Price sensitivity, bundle size, and negative correlation in tastes Our first test of the discriminatory theory examines the impact on the (basic) cable demand elasticity, ϵ_{bb}, of adding a network to a bundle. If bundling reduces household

Table 4 Estimates of the impact of the addition of cable networks on cable demand

Specification	OLS			IV: MSO Prices		
	Level effects	Slope effects	Elasticity effect	Level effects	Slope effects	Elasticity effect
General-interest networks						
WTBS	1.13	−0.02	−0.04	1.08	−0.02	0.03
	(0.20)	(0.01)	(0.06)	(0.22)	(0.02)	(0.08)
USA	0.41	−0.01	0.01	0.43	−0.01	0.03
	(0.16)	(0.01)	(0.05)	(0.18)	(0.01)	(0.07)
TNT	−0.04	0.00	−0.03	−0.11	0.00	−0.01
	(0.16)	(0.01)	(0.04)	(0.17)	(0.01)	(0.05)
Family	−0.30	0.01	0.03	−0.17	0.00	−0.01
	(0.18)	(0.01)	(0.07)	(0.19)	(0.01)	(0.08)
Nashville	0.29	−0.03	−0.12	0.29	−0.03	−0.13
	(0.15)	(0.01)	(0.04)	(0.16)	(0.01)	(0.06)
A&E	−0.11	−0.01	−0.04	−0.32	0.01	0.00
	(0.24)	(0.01)	(0.06)	(0.26)	(0.02)	(0.07)
General-interest average	0.23	−0.01	0.00	0.20	−0.01	0.00
	(0.07)	(0.00)	(0.00)	(0.08)	(0.01)	(0.00)
Special-interest networks						
Discovery	0.20	0.00	0.02	0.29	−0.01	0.00
	(0.13)	(0.01)	(0.05)	(0.14)	(0.01)	(0.06)
ESPN	0.40	−0.07	−0.22	0.49	−0.08	−0.28
	(0.26)	(0.02)	(0.03)	(0.31)	(0.03)	(0.05)
CSPAN	0.36	−0.02	−0.08	0.58	−0.03	−0.14
	(0.37)	(0.02)	(0.08)	(0.36)	(0.02)	(0.07)
Lifetime	0.33	−0.03	−0.10	0.40	−0.03	−0.13
	(0.19)	(0.01)	(0.05)	(0.20)	(0.01)	(0.06)
CNN	0.08	−0.02	−0.06	0.26	−0.03	−0.12
	(0.15)	(0.01)	(0.04)	(0.16)	(0.01)	(0.05)
Weather	0.18	−0.01	−0.03	0.27	−0.01	−0.06
	(0.16)	(0.01)	(0.03)	(0.17)	(0.01)	(0.04)
QVC	−0.27	0.00	0.00	−0.23	0.00	−0.03
	(0.34)	(0.02)	(0.06)	(0.37)	(0.02)	(0.07)
Learning	0.50	−0.04	−0.25	0.85	−0.07	−0.39
	(0.37)	(0.02)	(0.07)	(0.43)	(0.02)	(0.10)
MTV	0.11	−0.01	−0.04	−0.06	0.00	0.00
	(0.33)	(0.02)	(0.08)	(0.36)	(0.02)	(0.10)
Special-interest average	0.21	−0.02	−0.06	0.32	−0.03	−0.08
	(0.06)	(0.01)	(0.02)	0.08	0.01	0.02
Top-15 average	0.22	−0.02	−0.06	0.27	−0.02	−0.08
	(0.04)	(0.00)	(0.02)	(0.05)	(0.00)	(0.02)
α_b		0.01			−0.10	
		(0.10)			(0.14)	
α_I		0.07			−0.03	
		(0.10)			(0.13)	
α_{II}		−0.09			−0.31	
		(0.11)			(0.16)	
Price instruments	None			MSO prices		

Reported in first two of each group of columns are results from joint estimation of aggregate demand for basic service (b) and up to two expanded basic services (I, II). Number of observations is 1,159 for basic, 429 for expanded I, and 168 for expanded II. Reported network parameter estimates are constrained to be the same across services. Heteroscedasticity-consistent standard errors are reported in parentheses. Reported in third of each group of columns are the implied effect of adding that network on the basic demand elasticity (bootstrap standard errors in parentheses). Also reported is estimated price sensitivity for each service, $\hat{\alpha}_s$. Broadcast programming, demographic variables, and region dummies also included in all specifications. Specifications differ in their choice of instruments; see table bottom for instruments and Section 4.2.1 for instrument definitions.

preference heterogeneity, then one would expect $\partial\epsilon_{bb}/\partial X_c > 0, c = 1, \ldots, 15$. As above, estimates in bold represent statistically significant support for this hypothesis at size 0.05.[34]

The elasticity results provide qualified support for the theory. Combined, 10 of 15 networks are estimated to make cable demand more elastic, six of them significantly so. Furthermore, none of the five estimated positive effects is close to being statistically significant.

The breakdown of networks into groups according to the nature of their programming turns out to be informative. Among the six general-interest networks, only one (The Nashville Network) is estimated to significantly increase the demand elasticity, while five of the nine special-interest networks are estimated to do so. Indeed, the average impact to the basic demand elasticity of the 9 special-interest networks is −0.13 (0.03) while that of the 6 general-interest networks is −0.01 (0.02). The difference in these equals −0.11 (0.03), implying the hypothesis that the two effects are equal can easily be rejected. This conclusion—that it is the bundling of *special-interest networks* that increases the cable demand elasticity—is very robust, holding across a variety of model specifications, instruments, and time.[35]

Sensitivity to instruments Table 5 reports results of the demand model under alternative instrument sets. The results presented in the table differ only from those in Table 4 in their choice of instruments. The first three columns of Table 5 use measures of bargaining power as cost instruments. The last three (again) use MSO prices, but also instrument for the networks offered on each service with MSO Networks.[36]

Comparing the last three columns of Table 4 and the first three columns of Table 5, the point estimates differ little between the two sets of price instruments. As for MSO prices, instrumenting with the cost variables increases the estimated cable price sensitivity. Similarly, most networks have the expected level and sign effects as well as the expected impact on cable elasticities. Given the concern that MSO prices may not be appropriate instruments, the similarity of the results using different instrument sets for prices is encouraging.[37] As MSO prices vary considerably more than the cost shifters (cf. Appendix 2), the IV specification in Table 4 is chosen as the baseline specification for further analysis. The conclusions that follow, however, are robust to this choice.

[34]Specifically, one can reject the null hypothesis $H_0 : \partial\epsilon_{bb}/\partial X_k < 0$ v. $H_A : \partial\epsilon_{bb}/\partial X_c > 0$ at size 0.05.

[35]In particular, similar (unreported) regressions on more recent cable data from yielded similar results.

[36]See Section 4.2.1 for a detailed description of the instruments.

[37]Not reported are results using as instruments prices of other systems within the same MSO but outside the state or region of the given system. While consistent with the results presented here, the estimates are qualitatively similar, but less precise. For these reasons, we maintain a preference for the results presented above.

Table 5 Demand estimates with alternative instruments

Specification	IV: cost			IV: MSO prices/networks		
	Level effects	Slope effects	Elasticity effect	Level effects	Slope effects	Elasticity effect
General-interest networks						
WTBS	1.07	−0.02	−0.07	2.87	−0.11	−0.38
	(0.24)	(0.02)	(0.07)	(1.47)	(0.10)	(0.42)
USA	0.28	0.00	0.02	2.53	−0.09	−0.02
	(0.18)	(0.01)	(0.06)	(1.80)	(0.08)	(0.52)
TNT	−0.20	0.01	0.01	−0.48	−0.01	−0.20
	(0.17)	(0.01)	(0.05)	(1.78)	(0.10)	(0.31)
Family	−0.29	0.00	0.02	−0.35	0.10	1.61
	(0.19)	(0.01)	(0.06)	(1.81)	(0.12)	(0.97)
Nashville	0.19	−0.03	−0.08	0.02	−0.04	−0.29
	(0.17)	(0.01)	(0.05)	(1.36)	(0.08)	(0.22)
A&E	−0.41	0.01	0.03	0.19	−0.03	−0.22
	(0.28)	(0.02)	(0.07)	(1.96)	(0.11)	(0.60)
General-interest average	0.11	−0.01	−0.01	0.79	−0.03	0.08
	(0.09)	(0.01)	(0.02)	(0.55)	(0.03)	(0.30)
Special-interest networks						
Discovery	0.19	0.00	0.00	−0.14	0.00	−0.02
	(0.15)	(0.01)	(0.05)	(0.94)	(0.06)	(0.35)
ESPN	0.26	−0.06	−0.15	1.21	−0.12	−0.55
	(0.32)	(0.03)	(0.09)	(1.12)	(0.09)	(0.28)

CSPAN	0.74	−0.05	−0.17	2.86	−0.21	−1.37
	(0.36)	(0.02)	(0.06)	(1.79)	(0.12)	(0.60)
Lifetime	0.40	−0.03	−0.12	1.31	−0.10	−0.55
	(0.21)	(0.01)	(0.06)	(2.53)	(0.14)	(0.92)
CNN	0.18	−0.03	−0.08	0.57	−0.06	−0.30
	(0.16)	(0.01)	(0.04)	(0.92)	(0.06)	(0.29)
Weather	0.27	−0.02	−0.07	−0.41	0.08	0.50
	(0.17)	(0.01)	(0.03)	(1.16)	(0.09)	(0.69)
QVC	−0.24	0.00	−0.01	−0.68	−0.02	−0.30
	(0.39)	(0.02)	(0.08)	(1.59)	(0.09)	(0.46)
Learning	0.91	−0.07	−0.36	3.29	−0.15	−0.60
	(0.45)	(0.03)	(0.09)	(2.25)	(0.12)	(0.47)
MTV	0.12	−0.01	−0.05	−1.84	0.06	0.09
	(0.39)	(0.02)	(0.10)	(2.28)	(0.13)	(0.78)
Special-interest average	0.31	−0.03	−0.11	0.68	−0.06	−0.34
	(0.08)	(0.01)	(0.03)	(0.36)	(0.03)	(0.16)
Top-15 average	0.23	−0.02	−0.07	0.73	−0.05	−0.17
	(0.06)	(0.00)	(0.02)	(0.19)	(0.02)	(0.12)
α_{b}		−0.18			0.09	
		(0.14)			(0.36)	
α_{I}		−0.10			0.09	
		(0.14)			(0.32)	
α_{II}		−0.27			0.03	
		(0.15)			(0.30)	
Price instruments	Cost			MSO prices		
Network instruments	None			MSO networks		

Reported in first two of each group of columns are results from joint estimation of aggregate demand for basic service (b) and up to two expanded basic services (I, II). Number of observations is 1,159 for basic, 429 for expanded I, and 168 for expanded II. Reported network parameter estimates are constrained to be the same across services. Heteroscedasticity-consistent standard errors are reported in parentheses. Reported in third of each group of columns are the implied effect of adding that network on the basic demand elasticity (bootstrap standard errors in parentheses). Also reported is estimated price sensitivity for each service, $\hat{\alpha}_s$. Broadcast programming, demographic variables, and region dummies also included in all specifications. Specifications differ in their choice of instruments; see table bottom for instruments and Section 4.2.1 for instrument definitions.

The final group of columns, labelled "IV MSO Prices/MSO Networks," examines the sensitivity of these conclusions to the possible endogeneity of bundles themselves. To do so, we augment the instruments for prices with instruments for network carriage itself. As described earlier, these instruments are the average propensity to carry a given network on a given service for all other systems owned by the same MSO. The final group of columns presents the results of this specification when using MSO prices and MSO Networks as instruments. The results weaken somewhat. Twelve of the top-15 networks are now estimated to increase the basic cable price elasticity, but only 2 are statistically significant. This suggests network carriage within an MSO either proxies poorly for the costs of network carriage or does not cause sufficient variation in observed network carriage to identify the parameters cf. (cf. Angrist and Krueger 2001). As they do not provide strong evidence against the baseline specification, we maintain a preference for those results.

5.2 The consequences of bundling

What are the consequences of the heterogeneity reduction from bundling implied by these results? As we have taken a reduced-form approach in this paper, this is difficult to answer accurately.[38] It is worth having at least a rough estimate, however, both to evaluate the credibility of the results presented so far as well as to get some idea of the potential consequences of bundling for consumers and firms.

To obtain a rough estimate of the consequences of bundling, we implement a very simple exercise. We first calibrate our results to a simple model of bundle demand with normally distributed tastes. We then examine the impact of heterogeneity reduction of the order found in our results on firms' profit, consumers' surplus, and total surplus in the context of that simple model.

To calibrate our results to bundle demand with normally distributed tastes, we focus on the primary implication tested in this paper: that the bundle demand curve becomes more elastic with increases in bundle size. Recall our empirical results yield an estimate of the change in elasticity for each of the top-15 cable networks offered by systems (cf. column 6 in Table 4). We wish to equate these changes in elasticities with changes in the dispersion of preferences, σ_{bun}, in the model with normal tastes.[39] The advantage of this translation is that we can then use comparative statics on consumers surplus, profits, and total surplus drawn from the normal distribution to approximate the implied welfare effects of bundling (e.g. $\Delta\Pi \approx \frac{\partial\Pi}{\partial\sigma_{\text{bun}}} * \hat{\Delta}\sigma_{\text{bun}}$). These comparative statics results are given in Schmalensee (1984, p. S219).

[38]To accurately quantify the benefits of bundling, we would ideally calculate the effect on consumers, producers, and total welfare of alternative bundling strategies pursued by systems. To do so, however, requires estimates of the preferences for networks underlying the demand for cable bundles, namely their means, μ_j's, variances, σ_j^2's, and correlations, $\rho_{j,k}$'s. This is quite challenging and the topic of a separate work in progress (Crawford and Yurukoglu 2007).

[39]We focus exclusively on changes in preference heterogeneity via σ_{bun} as that is the primary consequence of bundling under the discriminatory theory.

To equate the reported changes in elasticity with those implied by changes in σ_{bun} we need estimates of marginal costs and the coefficient of variation, $\frac{\sigma_{\text{bun}}}{\mu_{\text{bun}}}$, in an "average" cable market. Following industry sources, we estimate marginal costs at 28% of current prices.[40] Estimating the "average" coefficient of variation is more difficult. To do so, we calculate the optimal price in a model with normal tastes for all possible values of $\frac{\sigma_{\text{bun}}}{\mu_{\text{bun}}}$ and then examine the predicted market share for the bundle at that price. Comparing the average market share of basic service in the data (68%) to those predictions yields an estimated coefficient of variation of 0.42.[41] At this value, the implied change in the basic demand elasticity per unit change in σ_{bun} equals 0.71. This allows us to convert our estimated changes in elasticities to changes in σ_{bun} by dividing each of the values in the third column of Table 6 by 0.71 (i.e. $\hat{\Delta}\sigma_{\text{bun},c} \approx \hat{\Delta}\epsilon_c / \frac{\partial \epsilon}{\partial \sigma_{\text{bun}}}$).

Table 6 presents the results of this exercise for the baseline specification (Specification (2) in Table 4). The first three columns of the table duplicate the estimated level, slope, and elasticity effects from the baseline specification. The next three columns report the associated percentage change in Consumers Surplus, Profit, and Total Surplus implied by each $\hat{\Delta}\sigma_{\text{bun},c}$.[42] Bootstrap standard errors are in parentheses.

Several interesting patterns emerge. As expected, consumer welfare falls from the heterogeneity reduction caused by bundling, with an average (across networks) per-network loss of 2.5% of their existing surplus. Similarly, firm profits rise due to the enhanced surplus extraction, with an average increase of 3.0% of their existing profits. Total surplus *increases*, with an average increase of 1.3%. These totals, however, mask important compositional effects across network types. General-interest networks, on average, do not increase the bundle demand elasticity and therefore do not influence profit or welfare (under the discriminatory theory).[43] Indeed, *(almost) all the profit and welfare effects are coming through the bundling of special-interest networks*. Bundling an average top-15 special-interest cable network is estimated to increase profits and reduce consumer welfare, with an average effect of 4.7% (4.0%) per special-interest network. On balance, total welfare increases, with an average effect of 2.0% per special-interest network.

The implications of these findings are quite interesting. As suggested by the sometime outrage over bundling in the industry, average consumer welfare from bundling of special-interest networks is estimated to fall. This suggests that consumers might benefit from having such networks available on an à-la-carte basis, even if the losses to firms outweigh their gains.

[40]See, e.g. Halfon (2003, footnote 78) and FCC (2003). While possibly high for the period we study, the qualitative conclusions we draw are similar across a range of marginal cost estimates.

[41]This suggests mean preferences for cable service bundles are, on average, approximately 2.40 times their standard deviation.

[42]Note that these results only quantify the profit and welfare consequences from the *heterogeneity reduction* caused by bundling. It ignores the (positive) effects to both profit and welfare of increases in the mean of preferences due to the addition of a valuable network to the bundle.

[43]An exception is the Nashville Network.

Table 6 Estimates of the impact of the addition of cable networks on elasticities and welfare

	IV: MSO Prices			Estimated welfare effects		
	Level effects	Slope effects	Elasticity effect	%Δ Profit	%Δ Cons. surplus	%Δ Total surplus
General-interest networks						
WTBS	1.14	−0.02	0.03	−0.020	0.017	−0.008
	(0.19)	(0.02)	(0.08)	(0.031)	(0.026)	(0.013)
USA	−0.19	−0.01	0.03	−0.012	0.010	−0.005
	(0.14)	(0.01)	(0.07)	(0.030)	(0.025)	(0.013)
TNT	0.28	−0.02	−0.01	0.000	0.000	0.000
	(0.15)	(0.01)	(0.05)	(0.020)	(0.016)	(0.008)
Family	0.02	−0.01	−0.01	0.016	−0.013	0.007
	(0.15)	(0.01)	(0.08)	(0.021)	(0.018)	(0.009)
Nashville	0.10	−0.01	−0.13	0.045	−0.038	0.019
	(0.13)	(0.01)	(0.06)	(0.022)	(0.018)	(0.009)
A&E	−0.33	0.01	0.00	0.000	0.000	0.000
	(0.23)	(0.01)	(0.07)	(0.016)	(0.013)	(0.007)
General-interest average	0.20	−0.01	−0.01	0.010	0.008	0.004
	(0.08)	(0.01)	(0.03)	(0.007)	(0.006)	(0.003)
Special-interest networks						
Discovery	0.07	−0.01	0.00	0.004	−0.003	0.002
	(0.12)	(0.01)	(0.06)	(0.021)	(0.018)	(0.009)
ESPN	0.58	−0.08	−0.28	0.094	−0.078	0.039
	(0.23)	(0.03)	(0.05)	(0.019)	(0.016)	(0.008)
CSPAN	0.47	−0.03	−0.14	0.053	−0.044	0.022
	(0.23)	(0.02)	(0.07)	(0.026)	(0.022)	(0.011)
Lifetime	0.28	−0.02	−0.13	0.042	−0.035	0.017
	(0.17)	(0.01)	(0.06)	(0.028)	(0.023)	(0.012)
CNN	0.32	−0.03	−0.12	0.052	−0.044	0.022
	(0.14)	(0.01)	(0.05)	(0.017)	(0.014)	(0.007)
Weather	−0.10	0.00	−0.06	0.021	−0.017	0.009
	(0.14)	(0.01)	(0.04)	(0.019)	(0.016)	(0.008)
QVC	0.09	−0.01	−0.03	0.006	−0.005	0.002
	(0.20)	(0.01)	(0.07)	(0.029)	(0.024)	(0.012)
Learning	0.71	−0.05	−0.39	0.144	−0.121	0.060
	(0.33)	(0.02)	(0.10)	(0.040)	(0.033)	(0.016)
MTV	−0.23	0.01	0.00	0.012	−0.010	0.005
	(0.25)	(0.01)	(0.10)	(0.047)	(0.039)	(0.019)
Special-interest average	0.32	−0.03	−0.11	0.047	−0.040	0.020
	(0.08)	(0.01)	(0.03)	(0.010)	(0.008)	(0.004)
Top-15 average	0.27	−0.02	−0.08	0.030	−0.025	0.013
	(0.05)	(0.00)	(0.02)	(0.006)	(0.005)	(0.003)

Reported are parameter estimates from the preferred specification (cf. Columns 4–6, Table 4) and associated estimated welfare effects. Welfare effects calibrated from a model of bundle demand with normally distributed tastes. Bootstrap standard errors in parentheses.

That being said, there are several caveats to these results. First, as noted above, the discriminatory effects are concentrated on the bundling of special-interest networks; bundling of general-interest networks is estimated to have no effect one way or the other on the bundle elasticity and thus on profit or welfare under the discriminatory theory. Indeed, it is just a handful of networks—ESPN, The Nashville Network, CNN, and The Learning Channel among them—that are driving our conclusions. Second, even our reported averages mask important distributional effects across consumers. The consumers that lose most from bundling are those that place high value on only one or a few networks in the bundle, but are still willing to purchase. For such consumers, bundling requires them to purchase unwanted channels, to the benefit of firms. By contrast, some consumers do gain from bundling. Bundling permits firms to lower prices (relative to the sum of unbundled prices) to the benefit of consumers that place moderate value on a large number of networks. Any structure, bundled or not, will have winners and losers among consumers.

Our final caveat is the most important. While the impact of bundling on the cable demand elasticities are robust, the profit and welfare effects reported above are from a simple calibrated model that does not fully capture all elements of the cable industry. In particular, we have assumed away the advertising market and any changes in programming, marketing, and administrative costs from offering networks on an unbundled basis. Relaxing any of these assumptions would tend to increase the affiliate fees (costs) to cable systems of unbundled offerings.[44] Crawford and Cullen (2007) and Crawford and Yurukoglu (2007) build on the ideas presented in this subsection to analyze the profit and welfare consequences of bundling under alternative assumptions on preferences and costs. Further discussion of these important issues is available there.

6 Conclusion

An influential theoretical literature supports a discriminatory explanation for product bundling: it reduces consumer heterogeneity, extracting surplus in a manner similar to second-degree price discrimination. While commonly advanced in the study of industrial organization, marketing, and business strategy, this is the first paper to explicitly test the implications of this theory and quantify its empirical relevance in the cable television industry.

The results provide qualified support for the discriminatory theory. For the preferred specification, carriage of 6 of the top-15 cable television networks is found to significantly increase the elasticity of the bundle (cable) demand curve (and never decreases it). Furthermore, as predicted by the theory, the effect of bundling on heterogeneity reduction is concentrated among special-interest networks, the preferences for which are likely to negatively co-vary in the population of consumers. Analysis of the implications of bundling in a simple model with normally distributed tastes calibrated to our results is suggestive of the empirical importance of these effects: bundling an average top-15 special-interest cable network yields an estimated reduction in consumer

[44] Kagan Media Research (2005) "estimates TV channel operators would need to raise per-capita channel carriage fees by a multiple of four to offset a 50% loss of subscribers from big basic bundles." Furthermore, Rennhoff and Serfes (2005) present a theoretical analysis of competition between program networks and cable systems with and without bundling and find costs to cable systems would increase without bundling.

heterogeneity equivalent to a 4.0% decrease in consumers surplus, 4.7% increase in firm profits, and 2.0% increase in total surplus.

An important implication of these findings is that the product choices of firms can be as important as prices in impacting consumer and social welfare. This nicely complements similar findings in the industrial organization literature of the welfare consequences of new goods (Griliches and Cockburn 1994; Bresnahan and Gordon 1996; Petrin 2003). It also highlights the importance of extending models of price competition widely used in merger and regulatory analysis to also consider firms' product choices (e.g. Einav 2007; Crawford and Shum 2006). Indeed, given the recent unbundling of elements of the local telephone, electric power, and software markets, assessing the benefits of extending competition and regulatory policymaking in this dimension is an important area of further research. Abstracting from any cost-side effects, the results presented here suggest there may be short-run social *losses* from unbundling which must be balanced against the gains from increased competition in components markets. Establishing empirical regularities of the competitive consequences of bundling is therefore of considerable interest.

Appendix

Appendix 1: Proofs of propositions

Proof of Proposition 1 [45] Suppose there are n discrete products (components) supplied by a monopolist and consumers differ in their preferences (willingness-to-pay) for each of these products, given by a type vector, $v_i = (v_{i1}, \ldots, v_{in})$. Let each v_{ic}, $c = 1, \ldots, n$, be independent with means μ_c and variances σ_c. Let $x_{in} \equiv \frac{1}{n}\sum_{c=1}^{n} v_{ic}$ be the per-good valuation for consumer i of a bundle of n goods, let μ_n be its mean, and let σ_n^2 be its variance. Note that μ_n and σ_n^2 follow the well-known formulas for the mean and variance of an average of (independent) random variables:

$$\mu_n = \frac{1}{n}\sum_{c=1}^{n} \mu_c = \bar{\mu}_c \tag{3}$$

and

$$\sigma_n^2 = \frac{1}{n^2}\sum_{c=1}^{n} \sigma_c^2 = \frac{1}{n}\left[\frac{1}{n}\sum_{c=1}^{n} \sigma_c^2\right] = \frac{1}{n}\bar{\sigma}_c^2 \tag{4}$$

Because the sequences v_{ic} are uniformly bounded, $\lim_{n\to\infty} \bar{\mu}_c$ and $\lim_{n\to\infty} \bar{\sigma}_c^2$ exist. Let $\lim_{n\to\infty} \bar{\mu}_c = \mu$ and $\lim_{n\to\infty} \bar{\sigma}_c^2 = \sigma^2$. Note this implies $\lim_{n\to\infty} \mu_n = \mu$ and $\lim_{n\to\infty} \sigma_n^2 = \lim_{n\to\infty} \frac{\sigma^2}{n} = 0$.

[45] Much of this proof follows ideas in Bakos and Brynjolfsson (1999, Appendix).

Let $q_n(p) \equiv \int_p^\infty dF(x_{in})$ give the market share of a bundle of size n offered at per-good price p, where $F(x_{in})$ is the CDF of x_{in}. Note that $q_n(\mu - \epsilon) = \text{Prob}(x_{in} > \mu - \epsilon) = \text{Prob}(x_{in} - \mu > -\epsilon)$.

Let $\epsilon^n(p) \equiv -\frac{\partial q^n(p)}{\partial p}\frac{p}{q_n}$ be the (absolute value of the) elasticity of the *per-good* demand curve evaluated at per-good price p and let $\tilde{\epsilon}^n(\tilde{p}) \equiv -\frac{\partial q^n(\tilde{p})}{\partial p}\frac{\tilde{p}}{q_n(\tilde{p})}$ be the corresponding (aggregate) elasticity of the bundle demand curve evaluate at total price $\tilde{p}$. For a bundle of size n, $\tilde{p} = np$.

By the weak law of large numbers and symmetry, $\text{Prob}(x_{in} - \mu < -\epsilon) \leq \frac{1}{2}\frac{\sigma^2}{\epsilon^2 n}$, implying $q_n(\mu - \epsilon) \geq 1 - \frac{1}{2}\frac{\sigma^2}{\epsilon^2 n}$. By a similar argument, $q_n(\mu + \epsilon) \leq \frac{1}{2}\frac{\sigma^2}{\epsilon^2 n}$.

Case I Per-good elasticity. We first prove the proposition for the per-good elasticity, ϵ^n.

Let $\omega = \frac{\sigma^2}{\epsilon^2}$ and consider a change in price from $\mu - \epsilon$ to $\mu + \epsilon$ on the per-good demand for a bundle of size n.

$$\begin{aligned} \frac{\partial q_n}{\partial p} &\leq \frac{1}{2}\frac{\omega}{n} - \left(1 - \frac{1}{2}\frac{\omega}{n}\right) \\ &\leq \frac{\omega}{n} - 1 \end{aligned} \tag{5}$$

Then

$$\begin{aligned} \epsilon^n(\mu - \epsilon) &= -\frac{\partial q_n}{\partial p}\frac{\mu - \epsilon}{q_n} \\ &\geq -\left(\frac{\omega}{n} - 1\right)\frac{\mu - \epsilon}{1 - \frac{1}{2}\frac{\omega}{n}} \\ &\geq \frac{n - \omega}{n - \frac{\omega}{2}}(\mu - \epsilon) \end{aligned} \tag{6}$$

Differentiating this with respect to the bundle size n yields

$$\frac{\partial \epsilon^n(\mu - \epsilon)}{\partial n} \geq \frac{\omega}{2\left[n - \frac{\omega}{2}\right]^2}(\mu - \epsilon) > 0 \tag{7}$$

Let $\epsilon = \mu - p_n^*$ so that $p_n^* = \mu - \epsilon$.[46] Then increasing bundle size makes the per-good bundle demand curve more elastic when evaluated at the profit-maximizing price for a bundle of size n.

Case II Aggregate bundle elasticity. When considering the impact of increases in n on the aggregate bundle elasticity, one has to accommodate that a given change in the per-good bundle price has a larger effect on the aggregate price for a larger bundle than for a smaller bundle. While this does not impact the elasticity of the size-n bundle demand curve evaluated at price p_n, it does

[46]Letting $\epsilon = \mu - p_n^*$ suggests we are looking at non-marginal changes in price. An alternative proof would show the elasticity of the bundle demand curve increases with bundle size when evaluate *at its mean*. Empirical implementation of that approach is more difficult, however, leading us to favor the given approach.

impact the elasticity of the size-$(n+1)$ bundle demand curve evaluated at price p_n. In particular,

$$\begin{aligned}\tilde{\epsilon}^{n+1}(\tilde{p}_n) &= \epsilon^{n+1}(p_n)\frac{n}{n+1}\\ &= \epsilon^{n+1}(p_n)A(n)\\ &\geq \frac{n-\omega A(n)}{n-\frac{\omega A(n)}{2}}\ \tilde{p}_n \end{aligned} \tag{8}$$

Under A4, it is easy to show that for the per-good elasticities, $\epsilon^{n+1}(p_n) \geq \epsilon^n(p_n)$. For the aggregate size-$(n+1)$ bundle elasticity, however, we must scale $\epsilon^{n+1}(p_n)$ by $A(n) < 1$. What impact does this have on the comparison? One can show that $\tilde{\epsilon}^{n+1}(\tilde{p}_n) \geq \tilde{\epsilon}^n(\tilde{p}_n)$ whenever the right-hand side of the last inequality in Eq. 8 is greater than the right-hand side of the last inequality in Eq. 6. This holds for all n.[47]

Proof of Proposition 2 Let preferences be as for Proposition 1 above except in allowing for correlation between consumer valuations, $v_i = (v_{i1}, \ldots, v_{in})$. Let $\rho_{c,d} = corr(v_{ic}, v_{id})$. With correlation, the variance of the per-good valuation for a bundle of size n, x_{in}, may be written as

$$\begin{aligned}\sigma_n^2 &= \frac{1}{n^2}\left(\sum_{c=1}^{n}\sigma_c^2 + 2\sum_{c=1}^{n-1}\sum_{d=c+1}^{n}\rho_{c,d}\sigma_c\sigma_d\right)\\ &= \frac{1}{n}\left[\frac{1}{n}\left(\sum_{c=1}^{n}\sigma_c^2 + 2\sum_{c=1}^{n-1}\sum_{d=c+1}^{n}\rho_{c,d}\sigma_c\sigma_d\right)\right]\\ &= \frac{1}{n}\bar{\sigma}_c^2 \end{aligned} \tag{9}$$

The primary benefit of bundling is due to heterogeneity reduction as measured by the variance of per-good tastes for the bundle, σ_n^2. Unlike for Proposition 1 above, once we allow for correlation in tastes, bundle size, n, is not a sufficient statistic for σ_n^2. In particular, adding a new good to a bundle changes σ_n^2 by both (1) changing $\bar{\sigma}_c^2$ and (2) increasing n.

Let $\eta = \frac{\omega}{n} = \frac{\sigma^2}{\epsilon^2 n} = \frac{lim_{n\to\infty}\bar{\sigma}_c^2}{\epsilon^2 n}$. Then we may re-write Eq. 6 above as

$$\begin{aligned}\epsilon^n(\mu-\epsilon) &= -\frac{\partial q_n}{\partial p}\frac{\mu-\epsilon}{q_n}\\ &\geq -(\eta-1)\frac{\mu-\epsilon}{1-\frac{1}{2}\eta}\\ &\geq \frac{1-\eta}{1-\frac{\eta}{2}}\ (\mu-\epsilon) \end{aligned} \tag{10}$$

[47] In particular, $\frac{n-\omega A(n)}{n-\frac{\omega A(n)}{2}}\ \tilde{p}_n \geq \frac{n-\omega}{n-\frac{\omega}{2}}\ \tilde{p}_n$ simplifies to $A(n) \leq 1$.

Differentiating this with respect to η yields

$$\frac{\partial \epsilon^n(\mu-\epsilon)}{\partial \eta} = \frac{-1}{2\left[1-\frac{\eta}{2}\right]^2}\,(\mu-\epsilon) < 0 \tag{11}$$

This is a more general statement of Eq. 7 above.[48] Reducing the (limiting) variance of the bundle (e.g. by increasing n or reducing σ^2) makes per-good demand for a bundle of size n more elastic.

The result of the proposition follows from Eq. 11. To see this, suppose the bundle had only two goods (i.e. component 2 was the nth good). It is easy to see that $\frac{\partial \sigma_n^2}{\partial \rho_{1,2}} > 0$, i.e. making the correlation between components 1 and 2 more negative reduces σ_n^2. Since, $\frac{\partial \eta}{\partial \sigma^2} > 0$ it follows that $\frac{\partial \epsilon^n}{\partial \rho_{\tilde{1},2}} < 0$: making correlations more negative makes the bundle demand curve more elastic. For the case of general n, simply note that the variance of a bundle of size n can be decomposed into the variance of a bundle of size $(n-1)$, the variance of component n, and twice the covariance between a bundle of size $(n-1)$ and component n.

$$\begin{aligned}
\bar{\sigma}_c^2 &= \frac{1}{n}\left(\sum_{c=1}^{n}\sigma_c^2 + 2\sum_{c=1}^{n-1}\sum_{d=c+1}^{n}\rho_{c,d}\sigma_c\sigma_d\right)\\
&= \frac{1}{n}\left[\left(\sum_{c=1}^{n-1}\sigma_c^2 + 2\sum_{c=1}^{n-2}\sum_{d=c+1}^{n-1}\rho_{c,d}\sigma_c\sigma_d\right) + \sigma_n^2 + 2\sum_{c=1}^{n-1}\rho_{c,n}\sigma_c\sigma_n\right]\\
&= \frac{1}{n}\left[\sigma^2_{n\tilde{-}1} + \sigma_n^2 + 2\rho_{n\tilde{-}1,n}\sigma_{n\tilde{-}1}\sigma_n\right]
\end{aligned} \tag{12}$$

Appendix 2: Instruments

In this appendix, we present an analysis of the instruments used for prices and network carriage in the econometric analysis.

Price instruments To assess the power of the price instruments, Table 7 presents results from reduced form regressions of prices on the instruments and exogenous variables.[49] The results are organized in sets of three columns. For each set of three, the first column reports the point estimates from the regression of the price of basic service, p_b, on the instruments and included exogenous variables. Similarly in the second and third columns for the price of Expanded basic services I and II, p_{I} and p_{II}, if offered.

[48] Note $\mathrm{sign}(\partial/\partial\eta) = -\mathrm{sign}(\partial/\partial n)$.

[49] Only results for the instruments are reported in the table.

Table 7 First-stage estimation, prices

Price inst: Cost				Price inst: MSO prices			
Instrument	Dependent variable			Instrument	Dependent variable		
	p_b	p_{I}	p_{II}		p_b	p_{I}	p_{II}
Homes passed, basic	0.00	0.00	0.00	IPB	0.64	0.03	0.13
	(0.03)	(0.01)	(0.01)		(0.07)	(0.10)	(0.20)
MSO subs, basic	0.01	0.00	−0.03	IPE	−0.03	0.34	−0.17
	(0.01)	(0.01)	(0.02)		(0.12)	(0.13)	(0.31)
MSO subs^2, basic	0.00	0.00	0.00	IPF	−0.22	0.48	−0.41
	(0.00)	(0.00)	(0.00)		(0.24)	(0.42)	(0.42)
Affiliation, basic	−3.48	0.96	0.36				
	(1.22)	(2.13)	(1.48)				
Channel capacity, basic	0.03	0.02	0.00				
	(0.08)	(0.02)	(0.00)				
Homes passed, expanded I	−0.02	–	–				
	(0.03)						
MSO subs, expanded I	0.00	–	–				
	(0.01)						
MSO subs^2, expanded I	0.00	–	–				
	(0.00)						
Affiliation, expanded I	6.12	–	–				
	(3.83)						
Channel capacity, expanded I	−0.05	–	–				
	(0.02)						
Homes passed, expanded II	0.01	−0.01	–				
	(0.01)	(0.01)					
MSO subs, expanded II	0.06	−0.04	–				
	(0.03)	(0.04)					
MSO subs^2, expanded II	0.00	0.00	–				
	(0.00)	(0.00)					
Affiliation, expanded II	−8.63	−1.54	–				
	(4.32)	(2.27)					
Channel capacity, expanded II	0.00	−0.01	–				
	(0.03)	(0.03)					
Observations	1,159	429	168	Observations	1,159	429	168
R-squared	0.649	0.818	0.888	R-square	0.720	0.818	0.888
p value	0.000	0.567	0.000	p value	0.000	0.029	0.350

Reported are results from reduced form estimation of prices for basic service (b) and up to two expanded basic services (I, II), on the instruments and exogenous variables. Results are organized in sets of three columns. The first set report estimates using Cost shifters as instruments, defined as homes passed, number and square of subscribers served by same firm (MSO), owner affiliation with programming networks, and channel capacity, interacted with cable service dummy variables. Separate effects for each service are not identified for some parameters in the Expanded Service equations. The second set of columns report estimates using MSO Prices as instruments, defined for each service as the average price for that service at other systems owned by the same MSO. Reported p value in each column is for hypothesis test of joint insignificance of reported parameters.

The first set of three columns report estimates using cost shifters as instruments for cable prices. As these shifters do not vary across services, we interact them with cable service dummy variables to allow their effects to differ

Table 8 First-stage estimation, network carriage

Instrument	Dependent Variable		
	NET_b	NET_I	NET_{II}
WTBS	<0.001	<0.001	0.002
Discovery	<0.001	<0.001	0.240
ESPN	<0.001	0.029	0.249
USA	<0.001	0.020	0.082
CSPAN	<0.001	0.229	–
TNT	<0.001	<0.001	0.065
Family	<0.001	<0.001	0.701
Nashville	<0.001	0.001	0.222
Lifetime	<0.001	<0.001	–
CNN	<0.001	<0.001	–
A&E	<0.001	0.016	–
Weather	<0.001	<0.001	0.067
QVC	<0.001	0.089	–
Learning	<0.001	0.084	–
MTV	<0.001	0.011	–
Other Satellite	<0.001	<0.001	<0.001
Observations	1,159	429	168

Reported are results of reduced form (probit) estimation of the carriage of each reported network on Basic Service (b) and up to two expanded basic services (I, II), on the instruments and exogenous variables. All specifications use MSO Networks as network instruments, defined for each network on each service as the proportion of other systems owned by the same MSO carrying that network on that service. Reported are p values for hypothesis test of joint insignificance of network instruments. Lack of carriage on Expanded Service II prevented identification of the impact of instruments for some networks.

by service. Reported are the estimated parameters for these interactions.[50] Evidence in support of the cost instruments is mixed. While homes passed does not appear to be an important cost shifter in any equation, the remaining variables enter intermittently. Most influential are affiliation (negative and significant in the first and third columns) and MSO subscribers and its square (negative for large values and occasionally significant in the first and third columns). Channel capacity enters as expected only in the second column. That said, p values associated with the hypothesis test of joint insignificance for all parameters are trivially small in all but the expanded I equation.[51] On balance, while supporting their use as instruments, lack of variation across services and an indirect connection to marginal costs suggests the cost shifters may be weak instruments.

[50]Note since all systems that offer an Expanded Service also offer a basic service, separate parameters are not identified for the Expanded I parameters in the second column. For an analogous reason, separate parameters are not identified for either Expanded Service in the third column.

[51]Note that because of the cross-equation restrictions on β and γ, identification obtains as long as the instruments are valid for at least one of the endogenous prices.

The second set of three columns report estimates using prices of cable services of other systems within an MSO as instruments.[52] The results are quite promising. Other-system prices within an MSO provide strong and significant effects for both basic and Expanded I equations, particularly for prices of the same service. Results for a second expanded service are poor, possibly due to relatively few observations. As expected, p values associated with the hypothesis of joint insignificance are soundly rejected for the basic and Expanded I equations.

Network instruments To assess the power of the network instruments, Table 8 presents a synopsis of reduced form (probit) regressions of network carriage on the instruments and included exogenous variables. As above, the results are organized in sets of three columns. As we must predict the carriage of each of the top-15 cable networks (as well as the sum of other cable networks) on all the exogenous variables and instruments, the number of estimations performed was considerable.[53] Rather than report the point estimates of the instruments for each specification, we simply report the p value from the hypothesis test of joint insignificance of the instrument set. As can be seen from the table, the instruments have considerable power, at least for the basic and first expanded basic equation.[54] Coefficient estimates were as expected—particularly powerful predictors of the carriage of network q on service s was the corresponding likelihood it was carried on service s by other systems within its MSO.

References

Adams, W., & Yellen, J. (1976). Commodity bundling and the burden of monopoly. *Quarterly Journal of Economics, 90*, 475–498.

Angrist, J., & Krueger, A. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. NBER Working Paper 8456.

Armstrong, M. (1996). Multiproduct non-linear pricing. *Econometrica, 64*(1), 51–75.

Armstrong, M. (1999). Price discrimination by a multi-product firm. *Review of Economic Studies, 66*, 151–168.

[52] Not all systems belong to an MSO. To address this issue, we pool systems with a single owner and treat them as other MSOs: including as a value for the instrument in market n the average price for all single-owner systems other than that in n. While the argument advocating marginal costs are correlated is weaker in this case than in the case of a common owner, single-owner systems tend to be smaller than average and, due to a common disadvantage when bargaining with network providers, have similar marginal costs.

[53] Specifically 16 networks * 3 services = 48 specifications.

[54] The problem of imprecision seen in the third column of the price regressions was exacerbated in the network carriage specifications. For some networks, only one or two of the 168 systems that offered a second expanded basic service carried that network on that service, implying an inability to identify the effects of any instruments. Analogously to that described above, identification of the parameters of interest requires the instruments have power for carriage on at least one service.

Bakos, Y., & Brynjolfsson, E. (1999). Bundling information goods: Pricing, profits, and efficiency. *Management Science, 45*(2), 1613–1630.
Bakos, Y., & Brynjolfsson, E. (2000). Bundling and competition on the Internet. *Marketing Science, 19*(1), 63–82.
Berry, S. (1994). Estimating discrete choice models of product differentiation. *Rand Journal of Economics, 25*(2), 242–262.
Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica, 63*, 841–890.
Bresnahan, T., & Gordon, R. (1996). The economics of new goods: An introduction. In T. Bresnahan, R. Gordon (Eds.), *The Economics of New Goods*. University of Chicago Press.
Carlton, D., & Perloff, J. (2001). *Modern Industrial Organization* (3rd ed.). Harper Collins
Chae, S. (1992). Bundling subscription TV channels: A case of natural bundling. *International Journal of Industrial Organization, 10*, 213–230.
Chipty, T. (1995). Horizontal integration for bargaining power: Evidence from the cable television industry. *Journal of Economics and Management Strategy, 4*(2), 375–97.
Chipty, T. (2001). Vertical integration, market foreclosure, and consumer welfare in the cable television industry. *American Economic Review, 91*(3), 428–453.
Crawford, G. (2000). The impact of the 1992 cable act on household demand and welfare. *RAND Journal of Economics, 31*(3), 422–449.
Crawford, G., & Cullen, J. (2007). Bundling, product choice, and efficiency: Should cable television networks be offered À La Carte? *Information Economics and Policy* (in press).
Crawford, G., & Shum, M. (2006). Empirical modeling of endogenous quality choice: The case of cable television. Mimeo, University of Arizona.
Crawford, G., & Yurukoglu, A. (2007). Demand, pricing, and bundling in subscription television markets. Mimeo, University of Arizona.
Einav, L. (2007). Seasonality in the U.S. motion picture industry. *RAND Journal of Economics, 38*(1).
FCC (2003). 2002 report on cable industry prices. Federal Communications Commission. Available at http://www.fcc.gov/mb/csrptpg.html.
FCC (2006). Further report on the packaging and sale of video programming to the public. Federal Communications Commission, February 2006. Available at http://www.fcc.gov/mb/csrptpg.html.
GAO (2003). Issues related to competition and subscriber rates in the cable television industry. Discussion paper, General Accounting Office, GAO-04-8.
Gentile, G. (2004). ESPN ends ugly fight with Cox over fees. *Associated Press*.
Griliches, Z., & Cockburn, I. (1994). Generics and new goods in pharmaceutical price indexes. *American Economic Review, 84*(5), 1213–1232.
Halfon, J. (2003). The failure of cable deregulation: A blueprint for creating a competitive, proconsumer cable television marketplace. United States Public Interest Research Group.
Hausman, J., Leonard, G., & Zona, D. (1994). Competitive analysis with differentiated products. *Annales d'Economie et de Statistique, 34*, 159–180.
Hazlett, T., & Spitzer, M. (1997). *Public Policy Towards Cable Television: The Economics of Rate Controls*. MIT Press.
Kagan World Media (1998). Economics of Basic Cable Television Networks. Discussion paper, Kagan World Media.
Kagan Media Research (2005). À la Carte pricing makes great theory, but tv channel bundling tough to beat. Kagan Media Research, December 2005. Available at http://www.ncta.com/IssueBrief.aspx?contentId=15.
Lancaster, K. (1971). *Consumer Demand: A New Approach*. Columbia University Press.
McAfee, R. P., McMillan, J., & Whinston, M. (1989). Multiproduct monopoly, commodity bundling, and correlation of values. *Quarterly Journal of Economics, 103*, 371–383.
Mitchener, B., & Kanter, J. (2004). EU member states back record fine for microsoft. *Wall Street Journal*, March 24, 2004.
Nalebuff, B. (2004). Bundling as an entry barrier. *Quarterly Journal of Economics, 119*(1), 159–187.
NCTA (1998). Cable television developments. Discussion paper, National Cable Television Association, Available at http://www.ncta.com.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica, 69*, 307–342.
Noam, E. M. (1985). Economics of scale in cable television: A multiproduct analysis. In E. M. Noam (Ed.), *Video Media Competition: Regulation, Economics, and Technology*. Columbia University Press.
Petrin, A. (2003). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy, 110*, 705–729.
Rennhoff, A., & Serfes, K. (2005). The role of upstream-downstream competition on bundling decisions: Should regulators force firms to unbundle? Mimeo, Drexel University.
Reuters (2003). US lawmaker urges a la Carte cable channel rates. *Reuters News Service*.
Rochet, J.-C., & Chone, P. (1998). Ironing, sweeping, and multidimensional screening. *Econometrica, 66*(4), 783–826.
Rochet, J.-C., & Stole, L. (2000). The economics of multidimensional screening. Mimeo, University of Chicago.
Salinger, M. (1995). A graphical analysis of bundling. *Journal of Business, 68*(1), 85–98.
Saloner, G., Shepard, A., & Podolny, J. (2001). *Strategic Management*. Wiley.
Schmalensee, R. (1984). Gaussian demand and commodity bundling. *Journal of Business, 62*, S211–S230.
Squeo, A. M., & Flint, J. (2004). Should cable be a la Carte not flat rate? *Wall Street Journal*.
Stigler, G. (1963). United States v. Loew's Inc.: A note on block booking. In P. Kurland (Ed.), *The Supreme Court Review: 1963* (pp. 152–157). University of Chicago Press.
Stigler, G. (1968). A note on block booking. In G. Stigler, R. D. Irwin (Eds.), *The Organization of Industry*.
Stole, L. (2003). Price discrimination in competitive environments. Mimeo, University of Chicago.
Thompson, A. (2006). NFL v. cable is turning into a real nailbiter. *Wall Street Journal*.
Varian, H. (2003). *Intermediate Microeconomics* (6th ed.). Norton
Waterman, D. H., & Weiss, A. A. (1996). The effects of vertical integration between cable television systems and pay cable networks. *Journal of Econometrics, 72*, 357–95.
Whinston, M. (1990). Tying, foreclosure, and exclusion. *American Economic Review, 80*(4), 837–859.
Wildman, S., & Owen, B. (1985). Program competition, diversity, and multichannel bundling in the New Video Industry. In E. M. Noam (Ed.) *Video Media Competition: Regulation, Economics, and Technology*. Columbia University Press.
Xiao, P., Chan, T., & Narasimhan, C. (2006). Product bundles under three-part tariffs. Mimeo, Washington University.