TESTING MODELS OF DIFFERENTIATED PRODUCTS MARKETS: CONSOLIDATION IN THE CABLE TV INDUSTRY*

BY DAVID P. BYRNE[1]

Research on differentiated products markets often uses structural demand/supply models to identify firms' marginal costs as product-level cost data is unavailable. Using unique demand and cost data from cable TV, I evaluate a differentiated products model's ability to identify marginal costs. I find firms systematically price below profit-maximizing levels, leading to biases in the model's marginal cost estimates. I study the implications for merger simulations, and find that these biases compromise estimates of merger-related cost efficiencies, yet do not prevent these models from generating useful predictions of the price and non-price effects of mergers.

JEL codes: L12, L96, C33, C52

Keywords: model validation, merger simulation, cable TV industry, consolidation, price regulation, endogenous product quality, discrete-choice demand, cost function estimation

Running head: Testing Differentiated Products Models

*Manuscript received June 2012, revised October 2013.

[1]This paper is based on a chapter from my 2011 Ph.D dissertation. I thank my advisors, Chris Ferrall and Susumu Imai, for their advice and support. The editor, Holger Sieg, and anonymous referees provided comments that substantially improved the paper. I have also received valuable feedback from Victor Aguirregabiria, Gaurab Aryal, Branko Bošković, Ingrid Burford, Mick Coelli, Chris Edmond, Matt Grennan, Joe Hirschberg, Sacha Kapoor, Stephen Law, Arvind Magesan, Simon Loerschter, Harry J. Paarsch, Carlos Serrano, Junichi Suzuki, Tom Wilkening, and seminar participants from Queen's, Simon Fraser, University of Alberta, HEC Montreal, Carnegie Mellon (Tepper), The University of Melbourne, University of Toronto, ANU, UNSW, UTS, Victoria University of Wellington, and Otago. Funding from the Social Sciences and Humanities Research Council of Canada and the University of Melbourne Faculty Research Grant scheme is appreciated. The article was previously circulated as "Consolidation and Price Discrimination in the Cable Television Industry" and "The Impact of Consolidation on Cable TV Prices and Product Quality." The online Appendix for this paper is available from the author's website. All errors and omissions are my own. Please address correspondence to: David P. Byrne, Department of Economics, The University of Melbourne, FBE Building, 111 Barry Street, Melbourne, VIC, 3010, Australia. E-mail: *byrned@unimelb.edu.au.*

1. INTRODUCTION

Empirical research on differentiated products markets is often hindered by missing information on firms' marginal costs. To overcome this problem, researchers often use structural demand and supply models to recover these costs. This typically involves two steps: (1) estimate a demand model with data on product-level prices, characteristics, and market shares; (2) "back out" the marginal costs implied by the first order conditions that govern firms' optimal pricing decisions, given assumptions on firms' conduct and costs. This technique has facilitated empirical research on various important issues in the *New Empirical Industrial Organization* such as measuring market power (Nevo 2001), quantifying the welfare gains from new products (Petrin 2002), identifying collusion (Bresnahan 1987), and evaluating mergers (Nevo 2000). Applications to merger evaluations have been particularly influential among anti-trust authorities who actively use merger simulations to assess the impact of prospective mergers.

Despite the prevalence and importance of this indirect approach to identifying marginal costs, there is little evidence on how well it works. There are at least two reasons for this. First, as mentioned, relevant information on costs is typically unavailable. By definition, differentiated products have different characteristics and therefore different costs of production. One would ideally want product-specific cost data to obtain marginal cost estimates that could serve as benchmarks in evaluating a structural model's ability to identify marginal costs. Unfortunately, such cost data tends to be highly sensitive and is generally not accessible to researchers.

The second reason relates to the identification of conduct. The implied marginal costs from a differentiated products model are only as good as a researcher's conduct assumption. For example, if one assumes all firms are Bertrand price competitors when in fact some coordinate on prices, then the marginal costs recovered from the Bertrand first order conditions will be biased. Thus, a researcher would also want to allow for an arbitrary form of conduct in recovering the structural model's implied marginal costs. However, as Nevo (1998) shows, identifying the conduct parameters of a differentiated products model is prohibitively difficult in practice because a large number of excluded demand instruments are required to do so.

This paper overcomes these difficulties to evaluate the performance of a differentiated products model in identifying marginal costs. I examine how conduct misspecification error gives rise to biased cost estimates, and study the implications for simulation-based merger evaluations. Through the analysis, I show how one can use cost data to improve upon the standard merger simulation approach in predicting the price, non-price, and welfare effects of mergers.

The paper uses data from the cable industry in Canada during the 1990s. For a number of reasons, this context is well-suited for such a study. First, uniquely detailed product-specific market-level cost data are available, in addition to demand-side data on prices, market shares, product characteristics, and demographics. Second, firms are licensed monopolists in pre-defined geographic markets during this period. From the outset, this should imply there is minimal uncertainty over conduct: firms aim to maximize profits within their local markets. Third, the data includes many (i.e., more than 100) mergers. This is partly due to the monopoly licensing scheme, which implies cable companies can only expand into new markets by acquiring other companies. This results in intense merger activity which peaks during the 1990s, a period of rapid cable consolidation both in Canada and around the world.[2]

Quantifying the welfare effects of cable consolidation may also be of independent interest. There has been surprisingly little empirical research on this topic despite the fact that cable mergers have long been of concern to authorities such as the Federal Communications Commission. This policy interest stems from the fact that cable television plays a major role in the leisurely activities for millions of households worldwide. Further, the industry has seen some of the largest scale mergers in history, such as the $60 billion AT&T-Comcast merger in 2002.[3]

The empirical analysis is developed over six sections. After describing the industry and data in Section 2, I develop and estimate a demand model for cable TV in Section 3. The model falls within the benchmark random coefficients logit discrete choice demand framework of Berry et al. (1995). Consumers have heterogeneous preferences for cable bundle quality and choose the cable package (basic or non-basic) that maximizes their utility.

The resulting demand-side estimates yield a number of results that reflect findings from previous research. The implied average own price elasticities for basic and non-basic of -4.21 and -5.45 are directly in line with previous estimates based on U.S. data. Moreover, I also find that nationally dominant cable companies offer higher per-channel programming quality than smaller firms. Based on the demand estimates alone, there is reason to believe that product quality (cable bundles) would have improved as a result of cable consolidation.

[2]Numerous authors have documented the history of consolidation in the cable industry. See, for example, Parsons (2003) for the U.S., Byrne (2010) for Canada, and Wieten et al. (2000) for the U.K. and Europe.

[3]See the FCC site http://transition.fcc.gov/transaction/attcocast.html for details of this merger, as well as other major mergers among media companies.

With the estimated demand model in hand, I turn to the supply-side of the model in Section 4. The model keeps with the Berry et al. (1995) environment by assuming that firms: (1) have constant marginal costs that are independent of the number of consumers in a market; and (2) maximize profits within their local markets. Two key features of the industry are also incorporated in the model. First, firms endogenously choose basic and non-basic cable prices, as well as the number of channels to include in each bundle.[4] Second, the price of basic cable is regulated. Unlike the U.S., Canada has a relatively simple, federal price regulation scheme. The regulations define a year-to-year market-specific basic cable price cap for markets with more than 2000 subscribers. Markets with fewer than 2000 subscribers do not face basic price regulation.

I validate the supply-side modeling assumptions by comparing the marginal costs inferred from the model's first order conditions to empirical marginal cost estimates based on the cost data. Two sets of empirical estimates are used for the investigation: average variable costs, and marginal costs based on an estimated translog cost function. The average-cost estimates equal marginal costs if marginal costs are indeed constant. The latter marginal cost estimates flexibly allow for increasing/constant/decreasing returns to scale. The cost function estimation results show the two sets of empirical marginal cost estimates are similar in magnitude. Auxiliary calculations using the translog cost function point to constant returns, which supports the constant marginal cost assumption typically used in the literature.

Validation results for the supply-side model show that the standard model does a poor job of identifying firms' costs. Specifically, the marginal costs implied by the model severely underpredict the marginal cost estimates based on the cost data. This may not be surprising in a regulated industry where pricing first order conditions may not hold. However, two additional findings suggest that the price regulation is not the sole driver of these results. First, I exploit cross-market variation in regulatory status to account for the impact that price regulation has on the model's implied marginal costs, and still find the implied marginal costs are far too small. Second, when I focus specifically on the unregulated markets in the sample, I find the implied marginal costs are even further from their empirical counterparts. Overall, the validation results imply firms are setting prices well below profit maximizing levels, which introduces conduct specification error into the structural model's cost estimates.[5]

[4]While the data on the number of channels in a given cable bundle are available, data on channel identities/bundle composition are not. As I discuss below, this forces me to adopt a slightly more complicated demand specification that allows for firm-specific differences in per-channel programming quality, possibly due to differences in bundle composition across firms.

[5]As I discuss in the next section, the time period for the study is prior to when cable companies started offering local or long-

Exploiting the richness of the cost data, I investigate whether differences between marginal revenue and costs systematically vary with firm or market characteristics. I refer to these differences as "conduct deviations." A number of interesting patterns emerge, most notably that nationally dominant firms exhibit much smaller conduct deviations than smaller firms. The results suggest that the supply-side of the benchmark differentiated products model better approximates the behavior of larger firms empirically.[6] There are numerous potential explanations for these results, one being that larger firms can afford to hire revenue managers to determine their profit-maximizing price levels, while small firms cannot.

Regardless of their underlying mechanism(s), the existence of systematic conduct deviations has important implications for the use of differentiated products models as empirical tools for informing policy. In Section 5, I study the implications for a particular policy-relevant application of these models: merger simulations.

An immediate implication of the validation results is that the implied marginal costs from the structural model cannot be used to quantify merger-related cost efficiencies. I highlight this by estimating and contrasting the marginal cost functions based on the cost data, and the implied costs from the model. For example, the data-based marginal cost functions clearly show that larger firms provide cable services at a lower cost per household. The marginal cost functions estimated with the implied costs are much noisier and do not reveal such cost-reducing firm size effects.

While the structural model is a poor measurement tool for identifying cost efficiencies (at least in the current context), it may still be useful for forecasting the price and non-price effects of mergers. To investigate this, I conduct a set of textbook/standard merger simulations. These simulations are based on a nonlinear system of equations that are derived from the standard differentiated products model and its implied marginal costs. In light of the supply-side model validation results, this nonlinear system cannot be strictly interpreted as corresponding to the first order conditions for profit maximization. This is because the implied marginal costs from the model are a function of the firms' true marginal costs and conduct deviations from profit maximization. Thus, the standard merger simulation approach incorporates differences in marginal

distance phone, Internet, or any other services. Therefore, the finding of systematic underpricing in cable cannot be a result of complementary product pricing for these products.

[6]This result reflects findings from Hortascu and Puller (2008). They find the bidding behavior of larger firms in the Texas electricity spot market closely align with theoretical predictions from static profit-maximization. In contrast, smaller firms are found to systematically use sub-optimal bidding strategies.

costs and conduct deviations across firms in predicting merger outcomes.

To highlight the importance of conduct deviations when predicting the price and non-price effects of mergers, I run a second set of modified merger simulations. I use the marginal cost estimates based on the cost data, but assume firms profit maximize subject to basic price regulation constraints. These simulations more accurately predict merger-related cost efficiencies. They have the potential disadvantage, however, of strictly relying on the profit maximization assumption in simulating merger outcomes. Its predictions over the price and non-price effects of mergers will be contaminated by deviations from this conduct assumption.

To evaluate the predictive accuracy of these two sets of merger simulations, I compare their simulated outcomes to difference-in-difference estimates of merger effects based on merger events in the data. This comparative structural and reduced-form approach to merger evaluation exploits the fact that the industry is consolidating. This yields many mergers in the sample with which to identify merger effects empirically.

The difference-in-difference estimates show that acquiring cable companies tend to reduce monthly basic cable prices by \$0.87, and add two more channels to the non-basic bundle. Non-basic prices and basic channel counts tend to remain unchanged. I further find that following acquisitions, dominant firms reduce monthly per-subscriber basic and non-basic non-content costs (labor and administrative expenses) by \$2.19 and \$1.63. Basic per-subscriber, per-channel programming costs fall by \$0.04, while non-basic programming costs tend to stay the same. I find the textbook merger simulations are able to reproduce the reduced-form estimates of the effects mergers have on basic cable prices and non-basic channels. Not surprisingly, however, they do a poor job at quantifying the merger-related cost efficiencies. In contrast, the modified simulations are able to closely predict the observed cost efficiencies, while their predictions over the impact of mergers on cable prices and channel counts are severely biased.

I further show how the standard merger simulation approach can be augmented with cost data to yield accurate predictions of the interrelated impact of mergers on cable prices, channel counts, *and* costs. This involves estimating marginal costs with the cost data, and using these estimates to separately identify the cost- and conduct-related components of the inferred marginal costs from the structural model. Using this integrated approach to simulating merger outcomes, I quantify the impact cable mergers have on consumer welfare and firms' profits. I find that cable mergers increase monthly per-subscriber consumer surplus and profits by \$2.39 and \$3.83, which are roughly 12% and 19% of average basic cable prices. I run auxiliary simulations to show that cost-reducing firm-size effects play an important role in generating these efficiency gains.

The paper's findings have various implications for policy and research on differentiated products markets. These are discussed at length in the conclusion in Section 6. The most immediate policy implication is that anti-trust authorities should use their power of subpoena to request cost data from firms when developing merger simulations. For example, the U.S. Hart-Scott-Rodino Act (1976) gives the Department of Justice and the Federal Trade Commission subpoena power to issue a "second request" for data from firms who are proposing to merge. My results motivate the use of such cost data in merger analyses since they can be used to improve the accuracy and transparency of the predictions from merger simulations.

1.1. *Related literature.* To my knowledge, this is the first paper to validate a differentiated products model using product-level cost data.[7] Previous studies by Nevo (2001) (ready-to-eat cereal) and Slade (2009) (beer) compare the inferred marginal costs from differentiated products models to aggregate industry-level marginal cost estimates. They conclude that market prices and costs more closely align with Bertrand price competition than pure collusion. However, because they do not have product-market level data, they are unable to quantify the degree to which conduct in individual markets depart from the Bertrand or any other conduct assumption. Relatedly, the authors are unable to study systematic biases in the model's predicted marginal costs under the assumption of Bertrand competition, nor the implications of these biases for measuring market power or evaluating mergers.

In contrast, my cost data permit a novel investigation of these systematic biases and the difficulties they create when estimating the model's cost parameters using conventional methods. In this way, my analysis is not unrelated to recent evaluations of the differentiated products model by Dube et al. (2012) and Knittel and Metaxoglou (2012), who study numerical difficulties in estimating the model's demand parameters.

This paper also complements previous studies that validate differentiated products models and merger simulations based on their ability to predict actual merger outcomes (Peters 2006; Weinberg and Hosken 2012; Houde 2012; Bjornerstedt and Verboven 2012). These papers collectively serve as a reply to Angrist

[7] A handful of papers have validated homogenous products models using product-level cost data: Genesove and Mullin (1998), Wolfram (1999), Clay and Troesken (2003), and Kim and Knittel (2003). In contrast to differentiated product markets, it is more reasonable to assume that all competitors face the same marginal costs in homogenous products markets. This greatly simplifies the problem of validating the supply-side of the model since one series of cost data can be applied to all products and firms in the industry. Moreover, Bresnahan (1982) shows it is straightforward to identify the single conduct parameter for homogenous products models, which further simplifies these model validation studies. These structural models have, however, received considerably less attention by policymakers compared to differentiated products models. This is especially true for empirical analyses of mergers.

and Pischke (2010), who call for additional retrospective merger analyses, and who raise concerns over the usefulness and reliability of structural models in IO. The current paper provides further evidence that addresses these issues on all fronts. The key difference is its focus on cost and conduct misspecification. In particular, I focus on how cost data can be used to improve the supply-side of the model, and sharpen the accuracy of and conclusions drawn from merger simulations. This differs from previous work where the emphasis is largely on the importance of demand specification for merger simulations.

Further, these and many other retrospective merger analyses assume exogenous product characteristics and focus solely on the impact of mergers on prices. This study considers both price and non-price (i.e., cable bundle quality) merger effects. Thus, this paper also relates to work by Draganska, Mazzeo, and Seim (2009) and Fan (2013) who incorporate endogenous prices and product characteristics in a differentiated products model to assess both the price and non-price effects of mergers. A key difficultly in these studies is that multiple Nash equilibria cannot be ruled out. This can compromise merger simulations since their models may have non-unique predictions over equilibrium merger outcomes.[8] These complications are not present in my context: firms are local monopolists, and I can easily find their unique profit-maximizing vector of prices and product characteristics.

Finally, the paper also adds to a large body of research on the cable TV industry. Recent papers by Crawford and Shum (2007), Chu (2010) and Crawford and Yurukoglu (2012) respectively study the impact of cable bundle quality degradation, Direct Broadcast Satellite (DBS) entry, and àla carte bundling in the cable TV industry. These papers use structural models that incorporate endogenous cable prices and product quality; these models are similar to the one I develop. This paper differs in its focus on cable consolidation which, like DBS entry, represents a significant change in industrial market structure and results in a major re-allocation of production toward more productive, dominant firms.[9] In this sense, the paper contributes to the broader empirical literatures on consolidation in media markets and industrial re-allocation.[10]

[8]To conduct merger simulations in such instances, one would ideally find every vector of equilibrium prices and product quality, and either specify an equilibrium selection rule, or report the non-unique predictions of the impact of mergers on prices, product quality, and welfare. Unfortunately, finding all equilibria for an arbitrary model specification is a prohibitively difficult problem. This prevents these authors from explicitly accounting for multiple equilibria in their merger simulations.

[9]A number of earlier descriptive papers from the U.S. document how cable prices and product characteristics vary with firm size and/or whether a cable company is vertically integrated with channel companies. Chipty (1995), Ford and Jackson (1997) and Chipty and Snyder (1999) all find evidence that horizontally integrated cable companies realize cost efficiencies.

[10]Various papers study consolidation in the U.S. radio industry following the 1996 *Telecommunications Act* and how it affects

2. INDUSTRY AND DATA

This section describes the context for the study: Canada's cable TV industry from 1990 to 1996. I first provide background on the industry during this period which, importantly, is prior to DBS entry (1998) and prior to cable companies offering telephone and Internet services (1999).[11] I then discuss my data sources and present some descriptive statistics. Finally, I document two facts: (1) larger companies offer cable packages with lower prices and more channels at a lower cost than smaller companies; and (2) the industry experiences consolidation that re-allocates national market shares toward a few dominant firms. These preliminary analyses highlight the many "moving parts" (cable bundle prices, channels, and costs) to consider when studying cable companies' behavior and the impact of consolidation. This motivates the use of a structural approach that models the joint determination of cable prices, channels, and costs within an internally consistent framework.

2.1. *Background and regulation.* Since 1968, cable companies in Canada have been federally regulated by the Canadian Radio-television and Telecommunications Commission (CRTC) according to the *Broadcasting Act* (the *Act*). Before 2001, a central feature of the *Act* was the issuance of geographic licenses that gave companies exclusive rights over the provision of cable services within pre-defined "Local Service Areas" (LSAs or licenses). License boundaries were defined by the CRTC and typically corresponded to cities, towns, or municipalities. In the period before DBS entry, these licenses effectively gave cable companies local monopolies within these pre-defined areas. These licenses were renewable, defined over three to five year horizons, did not involve fees, and could be revoked by the CRTC.

In the pre-satellite and pre-Internet era, cable companies primarily earned profits by offering tiered cable bundles in the form of basic cable (including the major broadcast networks like CBC, ABC, and NBC),

product variety (Berry and Waldfogel 2001; Sweeting 2010), impacts listener and advertiser welfare (Jeziorski 2013a; Mooney 2010), and generates cost efficiencies through economies of scale and scope (Smith and O'Gorman 2008; Jeziorski 2013b). The impact of consolidation on product repositioning has also been study in the context of the U.S. broadcast television (Calfee-Stahl 2011) and daily newspaper (Fan 2013) industries. More broadly, empirical studies of industrial consolidation fit within the re-allocation literature since they focus on a mechanism that yields a major re-allocation of resources towards acquiring firms. See Melitz and Redding (2014) for an overview of the literature on re-allocation in differentiated products markets, which in recent years has had a substantial impact on empirical research on international trade and productivity.

[11] Eastlink became the first cable company in Canada to offer telephone service in 1999, which signaled the start of a wave of "convergence" between traditional television and telephone service providers.

extended basic cable (including CNN, ESPN or TNT), and pay/specialty cable (including HBO or The Movie Network). The latter two tiers constituted "non-basic" or "discretionary" services that involved a tying requirement with basic cable. In particular, subscribers had to purchase basic cable as part of the non-basic package.

In most countries, the price and composition of cable packages were subject to some form of basic price regulation and channel carriage restrictions.[12] Relevant regulations are outlined in the 1986 *Cable Television Regulations*, which are a major amendment to the *Act*. Basic price regulation put an upper bound on the allowable increase in basic prices from year-to-year. This bound was largely determined by inflation: year-to-year nominal basic cable prices were allowed to increase up to 80% of the annual CPI-based inflation rate. Price increases were also permitted for cable licensees that were under financial distress, and to compensate companies that made major capital investments in cable systems.

Channel carriage restrictions involved two main components. First, they contained "must-carry" provisions that required companies to carry certain channels in their basic packages, including all local community programming, regional broadcast channels, and publicly funded educational and government proceedings channels.[13] In effect, these provisions created a minimum quality standard for each market's basic cable package, depending on the extent of local community broadcasting and the number of broadcast channels that were available.

Second, the CRTC licensed which channels were allowed to transmit their signals in Canada. Effectively, the CRTC maintained a list of channels that defined the universe of channels that cable companies in the country were allowed to offer in their basic and non-basic packages. Given this national list, companies were free to choose which channels they offered in their basic and non-basic cable packages, subject to the

[12] See Crawford (2013) for an extensive discussion of U.S. cable regulation. Market structure (i.e., local monopolies) and regulations over prices and channel carriage were similar in the Canada and the U.S. during the early-to-mid 1990s. A notable exception is that the U.S. experienced earlier DBS entry in 1994.

[13] Regional channels include any locally available television service operated by Canada's national public broadcaster, the Canadian Broadcasting Corporation. Other over-the-air national broadcast networks with regional affiliates, and who do not charge affiliation fees to cable companies, such as Global TV and the CTV Television Network, must also be carried in the basic cable package. Regional channels can also be province-specific, such as the French language broadcast channel TVA in Québec, or city-specific, such as CITY TV in Toronto. Other examples of channels falling under must-carry provisions include the Cable Public Affairs Channel (CPAC, which broadcasts proceedings from the Canadian House of Commons), educational services like TVOntario, or cultural broadcast channels like the Aboriginal Peoples Television Network (APTN).

must-carry restrictions just mentioned.

There were three regulatory classes that defined the degree to which the CRTC imposed these regulations: Class 1 (more than 6,000 subscribers), Class 2 (between 6,000 and 2,000 subscribers) and Part 3 (fewer than 2,000 subscribers). Class 1 and 2 licenses were subject to basic price regulation, with Class 1 licenses having more severe restrictions on allowable increases related to financial need or capital investments. Part 3 licenses did not face any basic price regulation in order to give cable companies an additional incentive to operate in small rural markets. All regulatory classes were subject to the carriage restrictions.

2.2. *Data and descriptive statistics.* The primary data sources are the CRTC Master Files from 1990 to 1996. These contain detailed annual information on basic and non-basic cable revenues, subscribership, and costs for each license in the country. These data are collected and verified by Statistics Canada and are used by the CRTC to monitor the cable industry. I further obtained descriptions of all variables included in the files using "Annual Return of Broadcasting Distribution Licensee" forms from the CRTC. These are the actual forms cable companies fill out when providing these license-level revenue and cost data to Statistics Canada each year. Importantly, the Master Files report the identity of the cable company that currently operates in each license/market in the country. For the remainder of the paper I will use "license" and "market" interchangeably.

On the demand side, the key variables include basic and non-basic subscription revenues, channel counts, and subscriber counts, as well as the total number of "homes passed" (i.e., the total number of people connected to a license's cable system). I calculate basic and non-basic cable prices by dividing subscription revenues by the number of subscribers for each tier.[14] Market shares are calculated as the total number of subscribers for a given cable package divided by the number of homes passed. Therefore, consumers of the "outside good" are households who are connected to the cable system, but who do not own TVs or who only watch free over-the-air channels using antennas.

On the supply-side, all basic and non-basic-related cable costs incurred by a cable company in each license and year are reported. These consist of the historical cost of capital assets and depreciation expenses, total salaries paid, sales and administrative expenses, and technical expenses. Capital and depreciation costs

[14]The data also contain the basic price charged by cable companies as well. Using this information, I have verified that my calculated basic prices correspond very closely to these reported prices. The Master Files do not distinguish between subscribership and revenues for the extended basic and specialty cable tiers. Therefore, my non-basic cable prices are the average revenue per-subscriber for extended basic and specialty cable.

correspond to expenses and investments for maintaining the cable system's head-end, distribution hub, and co-axial wiring, buildings, vehicles, and so on. Technical expenses correspond to cable companies' non-capital related costs of receiving and distributing programming content to their subscribers. These include affiliation payments made from downstream cable companies to upstream channel/content providers for offering their channels in basic or non-basic packages. Thus, the non-capital related components of costs can be grouped into two categories: *non-content costs* (salaries, sales and administrative expenses) and *content costs* (technical expenses).

Importantly, the Master Files also report whether a license falls under the Class 1, Class 2, or Part 3 regulatory scheme. The CRTC-approved maximum basic price a cable company can charge for a given license and year is also reported. These price caps are effectively determined by the basic cable price regulations described above.

After removing observations with missing data and dropping outliers, the resulting sample consists of an unbalanced panel of 5,415 license-year-product observations that span seven years across 3,102 licenses. So, for example, Montreal-1993-non-basic cable is an observation. Breaking the licenses down by regulatory class, there 403 Class 1, 686 Class 2, and 2,013 Part 3 licenses in the sample. Licenses can also be distinguished by the number of cable packages that are offered; there are 2,343 populous licenses with both basic and non-basic cable, and 759 (typically rural) licenses with only basic cable.

Table 1 presents sample means and standard deviations for the demand and cost variables; all dollar amounts are in 1992 constant dollars. The first column shows that subscribers in licenses with basic and non-basic cable on average pay \$18.76 for a basic package with 22 channels, and \$29.68 for a non-basic package with 22 basic channels and 10 non-basic channels.[15] On average, these licenses have 13,000 homes passed, with 43% and 42% of households subscribing to basic and non-basic cable.

[TABLE 1 ABOUT HERE]

To provide these services, cable companies typically incur monthly labor costs of \$3.74 and \$4.79 per basic and non-basic subscriber. They spend an additional \$4.89 and \$6.62 per basic and non-basic subscriber on sales on administrative costs each month.[16] Beyond these non-content costs, firms pay monthly basic and

[15]The maximum basic prices across all Class 1, Class 2 and Part 3 licenses are \$28.60, \$35.63, and \$43.34, respectively. The maximum basic price cables across all Class 1 and Class 2 licenses are \$28.60 and \$36.31.

[16]These are average monthly total labor costs for each cable package. For example, the \$4.79 non-basic labor cost includes the

non-basic content costs of \$6.36 and \$14.01 per subscriber. In per-subscriber, per-channel terms these costs are \$0.32 and \$2.30.[17] This cost differential provides some initial evidence that content providers for higher quality (i.e., more popular) non-basic channels are able to negotiate higher affiliation payments. Netting out all non-content and content costs from subscription prices yields average monthly per-subscriber profits of \$3.77 and \$4.26 for basic and non-basic cable. These correspond to average profit margins of 19% and 15%.

The fourth column of Table 1 presents analogous summary statistics for licenses with only basic cable. These licenses are considerably smaller, with only 1,200 homes passed on average. Relative to larger markets with basic and non-basic cable, these licenses tend to have higher cable prices, fewer basic channels, higher non-content and content costs, and much smaller profit margins of only 5%.

2.3. *Preliminary evidence of firm size effects on cable packages and costs.* Firm size, in terms of national market share, has a potentially large impact on a cable company's content costs, and a (related) indirect effect on its prices and cable bundles. Larger companies that are vertically integrated with upstream content providers can realize cost efficiencies that allow them to offer more channels and charge lower prices. Moreover, larger cable companies can negotiate lower affiliation payments made to non-integrated content providers. This bargaining advantage stems from the fact that content providers are concerned with the size of their audience; the larger the audience, the larger the payments they can command from TV advertisers. These payments are typically defined in terms of a fixed per-subscriber fee, and as such have a direct impact on the marginal cost of providing a basic and non-basic packages to consumers.[18]

The second and third columns of Table 1 provide preliminary evidence of such firm size effects. The second column lists summary statistics for the largest ten cable companies in terms of their national market share ("large firms"). The third column provides statistics for all other cable companies not in the top ten

\$3.74 basic labor cost. Thus, the average *incremental* labor cost per non-basic subscriber is \$1.05. Similar calculations apply for the sales and administrative costs and the content costs.

[17]The distribution of per-subscriber per-channel non-basic costs is right-skewed. Its median value is \$0.52.

[18]Various empirical studies from the U.S. have found evidence consistent cost-reducing firm size effects due to vertical integration or larger firms' bargaining advantage. These studies either provide indirect evidence based on observed relationships between firm size and cable bundle prices and composition (Ford and Jackson 1997; Chipty and Snyder 1999), or direct evidence that uses aggregate, average industry-wide affiliation payments for each channel in conjunction with a structural model of cable TV demand and supply (Crawford and Yurukoglu 2012).

("small firms").[19] Comparing the two columns, I find that on average large firms' basic and non-basic cable package prices are \$0.64 and \$3.18 lower, yet include one and five more channels, respectively. The average basic and non-basic market shares reflect these differences in cable packages: subscribers tend to choose non-basic cable in large firms' licenses, but not small firms' licenses.

Differences in large and small firms' costs are likely drivers of these differences in cable package prices and composition. This is evidenced by the average per-subscriber costs in columns two and three, which show that labor, sales and administrative, and content costs for basic and non-basic cable are all lower in larger firms' licenses. Among non-content costs, non-basic sales and administrative costs exhibit the largest difference; on average these costs are \$2.90 (or 37%) lower in large firms' licenses. Since non-content costs are local to a cable system, they are likely not affected by differences in vertical relations between content providers and large and small firms. One interpretation of these cost differentials is that larger firms are generally more experienced in managing cable systems which, all else equal, yields efficiencies in labor, sales, and administrative costs.

Among content costs, the differences in non-basic costs are most pronounced: large firms' average monthly per-subscriber content costs are \$3.79 (or 24%) lower than that of small firms. This cost difference is amplified on a per-subscriber, per-channel basis, where large firms' costs per channel and subscriber are more than 50% lower than small firms' costs (i.e., \$1.50 versus \$3.53).

Overall, these patterns are consistent with a story where dominant cable companies realize cost efficiencies that allow them to offer richer cable bundles and charge lower prices.[20] The profit figures in the table show this ultimately translates into substantive profit differentials between large and small firms. On average, large firms realize basic and non-basic margins of 31% and 27%, which compare to margins of only 9% and 6% for small firms.

[19]These firm size rankings are based on national market share are in terms of the number of basic or non-basic cable subscribers served by a cable company in 1993, the middle year of the sample. The firm-specific national market share measures are based on the universe of more than 1,200 cable licenses in the country, as described in Byrne (2010). The rankings are largely unchanged across years, and the results are the same if I classify "large firms" as those in the top 5, 10, or 20 in terms of national market share.

[20]Systematic differences in the characteristics of cable licenses operated by large and small firms could simultaneously explain these differences. For example, recall from Table 1 that larger firms tend to operate in more populous markets than small firms. To the extent that local wages in larger markets are lower than those in rural markets (possibly because of differences in local labor market competition), the average labor costs of large firms will be systematically lower than those of small firms. In the econometric models below, I control for these license-specific effects in identifying firm size effects in non-content and content costs.

2.4. *Industrial consolidation.* The large profit differentials between large and small firms, combined with the local monopoly regulation, suggests that large cable companies had a strong incentive to acquire small companies during the 1990s.[21] These incentives were complemented by the CRTC's laissez-faire approach to cable mergers. The regulator evaluated prospective mergers on a case-by-case basis, and by and large rubber stamped proposals that promised cost efficiencies and improved cable TV services for subscribers affected by a merger. Nationally dominant cable companies often made such claims in proposing acquisitions.[22] Technological change further promoted cable consolidation, both in Canada and around the world. In particular, the emergence of new fibre-optic technologies, and the expectation of telecommunications convergence among cable and telephone providers, gave dominant cable companies further incentives to expand and establish market presence.

Figure 1 illustrates the consolidation process that re-shapes the industry from 1985 to 2002.[23] The left panel plots the total number of cable mergers each year, as well as the number of mergers that involve acquisitions by the largest five and ten companies (by national market share). Though there are 300 to 400 cable companies operating in a given year, a relatively large proportion of acquisitions (48%) involve one of the ten dominant firms purchasing a small firm. In total, there are 449 mergers over this period

[21] As I document in Byrne (2010), there was active entry into new, previously unserved cable TV markets by incumbent cable companies in the 1980s. These new markets were, however, largely rural areas; by 1990 the vast majority of households in Canada had access to cable TV. After this point, cable companies mainly expanding into new markets through acquisition.

[22] More specifically, the CRTC put the onus on the parties involved to show that a proposed acquisition "yields significant and unequivocal benefits to the communities served." I provide an example of a CRTC-documented merger decision that involves improved channel offerings in online Appendix A.

[23] These figures are based on an auxiliary dataset that is described in Byrne (2010). In short, I construct these series by tracking merger events, firm identities, and subscribership each year for the universe of more than 1,200 cable licenses in the country. I identify the timing of cable mergers and accurately track a given license's cable company year-to-year using the CRTC Decisions, Notices, and Orders (DNO) Files from 1985 to 2002. For each year, they provide detailed reports of any cable mergers and the licenses they affect. They are available online at http://www.crtc.gc.ca/eng/dno.htm; in online Appendix A I provide an example DNO file for a cable merger. I use the accurate firm identifiers for each license and year that emerge from my analysis of the DNO files to construct the sample used in this paper. The subscribership figures are based the total number of subscribers reported for each license in the Master Files from 1985 to 1996 (which unfortunately do not differentiate between basic and non-basic counts from 1985 to 1989). Using these data, as well as information on license-level household counts from Statistics Canada from the 1986, 1991, 1996, and 2001 Censuses, I impute any missing data on license-level subscribership from 1985 to 1996, and extrapolate the total number of subscribers for each license from 1997 to 2002.

(26 mergers per year), reaching a peak of 49 mergers in 1990. A merger between two companies sees the acquiring firm take over 3 licenses on average. However, there is a lot of heterogeneity in the size of mergers; 67% of acquisitions see only one license get transferred from one firm to another, but there are also major acquisitions such as Rogers's purchase of MacLean-Hunter in 1994 which saw 33 licenses change hands.

[FIGURE 1 ABOUT HERE]

The right panel of Figure 1 illustrates the impact that consolidation has on national market shares for the dominant firms. The five largest companies see their collective national market share grow from 45% in 1986 to 84% in 2002. Table 2 further breaks down of national market shares and the number of licenses run by the ten largest firms. The table highlights how Rogers, Shaw, Vidéotron, and Cogeco come to dominate the industry in the 2000s. The dominance is largely regional with Shaw mainly owning licenses in Western Canada, while the others operate in the East. Their market shares partly reflect the fact they operate in the country's largest markets such as Montreal (Vidéotron), Toronto (Rogers), and Vancouver (Shaw). Indeed, while the largest ten companies served 84% of all subscribers in 2001, they only ran 411 (33%) of the licenses in the country. Thus, there is a non-negligible number of rural licenses operated by one of many local cable providers. These include proprietors who service their local market, and community cooperatives that jointly fund a local cable system.

[TABLE 2 ABOUT HERE]

Given the preceding set of facts, it is natural to ask, is what impact does consolidation have on cable prices, channels, and costs? I return to this question in Section 5 when I present and contrast reduced-form and structural estimates of merger effects.

3. DEMAND

This section develops and estimates a demand model for the cable television industry. I first describe a model where consumers have heterogeneous preferences for cable, and choose the package that maximizes their utility. I then discuss how I estimate and identify the model's parameters, and close by presenting the results.

3.1. *Model.* The unit of analysis is a license-year. Cable packages are indexed by j, where consumers choose whether to purchase basic cable ($j = 1$), non-basic cable ($j = 2$), or not to purchase cable at all

($j = 0$, the outside good). The number of products offered in license m at time t is denoted J_{mt}; markets with only basic cable have $J_{mt} = 1$ products, while markets with basic and non-basic cable have $J_{mt} = 2$ products.

Consumer i's indirect utility from buying package j in license m and year t is specified as:

$$
\begin{aligned}
u_{ijmt} &= t_i q_{jmt} - \alpha p_{jmt} + y^0_{jmt}\psi_0 + \xi_{jmt} + \varepsilon_{ijmt} \\
&= t_i q_{jmt} + \delta_{jmt} + \varepsilon_{ijmt} \qquad (1)
\end{aligned}
$$

where $\delta_{jmt} \equiv -\alpha p_{jmt} + y^0_{jmt}\psi_0 + \xi_{jmt}$ is the "mean utility" that is common to all consumers for product j in market m in year t. Under this specification, consumers i's utility depends on bundle j's price (p_{jmt}), programming content quality (q_{jmt}), non-content related factors (y^0_{jmt}, ξ_{jmt}), and consumer i's preference shocks (t_i, ε_{ijmt}). I now describe the latter three non-price components of the utility function in turn.

3.1.1. *Content quality.* The quality of progamming content in bundle j is specified as follows:

$$q_{1mt} = (z_{mt}\beta_1) \cdot x_{1mt} \qquad (2a)$$

$$q_{2mt} = (z_{mt}\beta_1) \cdot x_{1mt} + (z_{mt}\beta_2) \cdot x_{2mt} \qquad (2b)$$

where x_{jmt} is the number of channels in bundle j in license m and year t. The additive specification for q_{2mt} reflects the tying constraint that requires consumers to purchase basic cable as part of the non-basic package. This requires that non-basic cable has higher prices and bundle quality than basic cable: $q_{2mt} > q_{1mt}$ and $p_{2mt} > p_{1mt}$.

The variables in z_{mt} affect the rate at which bundle quality grows with additional channels. These include a multi-system operator (MSO) dummy variable (MSO_{mt}) that equals one if license m's cable provider in year t operates in more than one license across the country. The vector also includes firm-specific dummy variables for the largest 4 firms in the sample: Rogers, Shaw, Cogeco, and Vidéotron ($LargeFirm_{mt}$). Thus, the parameters in β_1 and β_2 capture differences in the marginal utility from additional basic and non-basic channels among dominant companies, smaller MSOs, and single system cable operators.[24]

I expect the β_1 coefficients to be small relative to the β_2 coefficients. Basic channels largely consist of local programming and national broadcast networks, channels that historically have been offered in the

[24] So, for example, if $\beta_2 > 0$ for Rogers, then a non-basic bundle of ten channels offered by Rogers will be of higher quality than a non-basic bundle of ten channels offered by a small locally-owned cable operator.

basic packages of both national cable companies and locally-owned companies (which causes β_1 to be small). Moreover, national cable providers historically offered more diverse non-basic packages than small companies (which causes β_2 to be large).

The q_{jmt} index is an aggregate content quality measure in the sense that the individual qualities of the channels included in bundle j ultimately determine its overall quality. This specification of content quality reflects the main limitation of my data, namely that data on license- or firm-specific channel identities are not available. This prevents me from estimating channel-specific effects on bundle quality as previous authors have done with U.S. data on channel identities (Crawford 2000; Crawford and Yurukoglu 2012). In the absence of data on channel identities, I choose to exploit my channel count data and the richness of my panel (in particular, that dominant cable companies operate many licenses) to construct a quality index that allows bundle size and firm-specific differences in per-channel quality to determine overall bundle quality.[25]

Beyond my lack of data on channel identities, a secondary motive for using an aggregate quality index is that it greatly simplifies the supply-side of the model. As I describe in Section 4 below, I assume firms choose the prices and aggregate qualities of their basic and non-basic cable packages to maximize profits within each of their licenses. An alternative and perhaps more realistic approach would be to model the optimal channel bundling choices of firms. Unfortunately, this problem suffers from a Curse of Dimensionality since it is extremely difficult to find the profit-maximizing subset of channels from the power set of all possible channel combinations[26] Modeling this process would therefore introduce a severe computational burden on

[25]I have also explored an alternative approach where I treat quality as an unobservable, and infer the values of q_{jmt} that rationalize observed cable prices and license shares (see Byrne 2010 and Crawford and Shum 2006). However, such an approach significantly complicates the use of the Berry, Levinsohn, and Pakes (1995) share inversion algorithm to recover unobserved non-content cable quality ξ_{jmt} (discussed in Section 3.2 below). If one treats q_{jmt} and ξ_{jmt} as unobserved to the econometrician, then for any value of q_{jmt} one can find a ξ_{jmt} that equates the model's predicted license shares for bundle j in license m at time t to its empirical counterpart. This implies that ξ_{jmt} cannot be recovered using the share inversion algorithm alone if q_{jmt} is treated as an unobservable. Recovering this unobserved heterogeneity is necessary for addressing the first-order concern of the impact unobserved non-content quality has on consumer demand, and the pricing and bundle quality decisions of companies. Allowing for this heterogeneity is critical for both estimation and in conducting counterfactual simulations.

[26]To get a sense of the size of the problem, consider a typical bundling problem for a cable company in a market with basic and non-basic cable. Recall that on average, 19 and 10 basic and non-basic channels are offered in these markets. For the sake of illustration, assume these channels are respectively chosen from a larger set of 25 and 15 basic and non-basic channels that the firm considers offering to its subscribers (from the CRTC's DNO decision files these figures are potentially much larger than this). Then to find the optimal cable package to offer, the firm would have to compute the optimal prices and profits from $2^{25} \times 2^{15} =$

my merger simulations in Section 5. I avoid these difficulties by using an aggregate quality index. Chu, Leslie, and Sorensen (2011) find that simple pricing rules based on bundle size (i.e., the number of channels in a package) is a good approximation to firms' optimal bundling decisions, which lends support to this modeling simplification.

This approach to modeling bundle quality does, however, introduce some important limitations for my analysis. The demand parameters β_1 and β_2 are reduced-form since they are a function of firms' underlying optimal bundling decisions. These firm-specific bundle quality coefficients should therefore be interpreted as short-run differentials in per-channel content quality among dominant firms for the 1990-1996 period, where regulation, technology, and optimal bundling decisions across firms are largely unchanged. These parameters would likely evolve over longer time horizons due to technological change or new channel entry into the market.

In addition to limiting the interpretability these demand parameters, abstracting from the bundling problem also hinders the interpretation of some supply-side parameters and results regarding the impact of consolidation. I defer discussion of these issues until I have described firms' profit-maximization problem in Section 4 and merger simulations in Section 5.

3.1.2. *Non-content related factors.* The vector y^0_{jmt} contains the variables that affect consumers' preferences for cable that are not related to bundle prices nor quality. These variables can be interpreted as factors that affect non-content quality. These include the MSO dummy, dummy variables for individual MSOs in the sample ($Firm_{mt}$), a non-basic cable dummy variable that equals one if bundle j is non-basic cable ($NonBasic_{jmt}$), a vector of six year dummies ($Year_t$), and a vector of eight province dummies ($Prov_m$). The firm dummy variables capture differences in non-content quality across small single-system cable companies, nationally dominant firms, and non-dominant multi-system operators, possibly related to firm-specific differences in branding or service quality. The non-basic dummy variable captures any non-content related differences in demand between basic and non-basic cable. The year and province dummies account for differences in cable demand over time and across the country.

Importantly, I allow for unobserved heterogeneity in non-content quality, ξ_{jmt}, that both consumers and firms account for in their decision-making, but which is unobserved to the econometrician. I normalize all deterministic parts of the utility function for the outside option to 0 since I do not observe its characteristics

$33,587,200$ possible bundle combinations.

(such as the number of over-the-air channels or signal strength). Under this normalization, ξ_{0mt} accounts for the quality of the outside good in license m and year t.

3.1.3. *Preference shocks.* Consumers' vertical taste coefficients for cable quality, t_i, are i.i.d draws from a Weibull distribution with license-specific scale and shape parameters ρ_{mt} and κ_{mt}. The Weibull distribution is attractive because it can flexibly model many single-peaked distributions. The Weibull assumption also ensures that individuals' marginal utilities for cable quality are non-negative. By allowing these taste parameters to vary with market characteristics, I admit a great deal of flexibility in substitution patterns between basic and non-basic cable across different licenses.

Following Chu (2010), I restrict ρ_{mt} and κ_{mt} to reasonable ranges by assuming the following functional forms:[27]

$$\rho_{mt} = \exp(y^1_{mt}\psi_1) \tag{3a}$$

$$\kappa_{mt} = 0.1 + 14.9 \cdot \left(\frac{\exp(y^2_{mt}\psi_2)}{1+\exp(y^2_{mt}\psi_2)}\right) \tag{3b}$$

The variables in y^1_{mt} that affect the *scale* of the vertical type distribution include license-level measures of average household income (Inc_{mt}), urban density (Urb_{mt}), and a dummy that equals one if license m in year t has both basic and non-basic cable ($TwoPackages_{mt}$). The variables in y^2_{mt} that affect the *shape* of vertical taste shock distribution include license-level total population (Pop_{mt}), the variance in household income within a license ($IncVar_{mt}$), and the $TwoPackages_{mt}$ dummy variable. The demographics exogenously affect the level and variability of cable demand. I expect licenses with higher average household income and lower urban density to have larger demand for cable. Demand for cable TV should be higher in more rural markets because consumers in these markets have fewer leisurely substitutes to watching cable. Larger licenses and licenses with more income variability are expected to have more heterogeneity in individuals' preferences for cable quality. I include $TwoPackages_{mt}$ in y^1_{mt} and y^2_{mt} to account for differences in tastes across licenses with just basic cable, and licenses with both basic and non-basic cable.

Consumer i also experiences an idiosyncratic horizontal preference shock for bundle j, ε_{ijmt}. I assume these shocks are i.i.d draws from a mean zero Type 1 Extreme Value distribution with scale parameter

[27] These functional forms restrict ρ_{mt} to be positive and κ_{mt} to lie in the large range of 0.1 to 15. The scale of the distribution, $1/\rho_{mt}$ is inversely proportional to ρ_{mt}.

normalized to $\sigma = 1$.

3.1.4. *Market shares.* Consumers choose the package from their market's menu of basic and non-basic cable packages $\{(p_{jmt}, q_{jmt})\}_{j=1}^{J_{mt}}$ that maximizes their indirect utility. Conditional on an individual's vertical type t_i, the market share for bundle j can be computed directly using the familiar logit formula:

$$s_{jmt}(\mathbf{p}_{mt}, \mathbf{x}_{mt}; t_i) = \frac{\exp(t_i q_{jmt} + \delta_{jmt})}{1 + \sum_{j'=1}^{J_{mt}} \exp(t_i q_{j'mt} + \delta_{j'mt})} \tag{4}$$

where $\mathbf{p}_{mt}$ and $\mathbf{x}_{mt}$ are $J_{mt} \times 1$ vectors that stack prices and channel counts for packages $j = 1 \ldots J_{mt}$. The license share for bundle j can be computed by integrating (4) over the vertical type distribution that is specific to market m and year t:

$$s_{jmt}(\mathbf{p}_{mt}, \mathbf{x}_{mt}; \theta) = \int s_{jmt}(\mathbf{p}_{mt}, \mathbf{x}_{mt}; t_i) f(t_i; \rho_{mt}, \kappa_{mt}) dt_i \tag{5}$$

where $f(\cdot; \rho_{mt}, \kappa_{mt})$ is the probability density function for the Weibull distribution, conditional on scale and shape parameters ρ_{mt} and κ_{mt}, and $\theta = \{\alpha, \beta_1, \beta_2, \psi_0, \psi_1, \psi_2\}$ contains all the demand parameters. Denoting Q_{mt} at the exogenously given number of homes passed by the cable system in license m and year, the demand for bundle j is the share times the number of homes passed: $Q_{jmt} = s_{jmt} Q_{mt}$.

3.2. *Estimation.* I estimate the demand parameters using the nonlinear GMM approach of Berry, Levinsohn, and Pakes (1995), where I assume the following population moments hold at the true value of the demand parameters, θ_0:

$$E[\xi_{jmt}(\theta_0) | \mathbf{Z}_{jmt}] = 0$$

where $\mathbf{Z}_{jmt}$ is the vector of instruments. To recover the demand shock $\xi_{jmt}(\theta)$ for a given θ, I first use the contraction mapping approach of Berry, Levinsohn, and Pakes (1995) to find the vector of mean utilities in market m and year t, $\delta_{mt}(\theta)$ that equates the predicted market shares of the model to those observed in the data. After finding the mean utilities, I can recover the unobserved non-content product quality, $\xi_{jmt}(\theta) = \delta_{jmt}(\theta) + \alpha p_{jmt} - y_{jmt}^0 \psi_0$.

The set of instruments include all the exogenous demand variables[28] as well as eight instruments that are excluded from the demand model: the size of a cable company in terms of the total number of subscribers

[28] These variables are: $NonBasic_{jmt}, MSO_{mt}, Firm_{mt}, Year_t, Prov_m, Inc_{mt}, Urb_{mt}, Pop_{mt}, IncVar_{mt}, TwoPackages_{mt}$.

it served nationally in the previous year ($FirmSize_{mt}$); a cable system's total potential size ($SysSize_{mt}$); channel capacity ($ChanCap_{mt}$); total kilometers of cable used ($KmCab_{mt}$);[29] regulatory dummies for Class 2 and Part 3 licenses ($Class2_{mt}$ and $Part3_{mt}$); and the average price and total number of channels offered on cable bundle j across all other cable systems in market m's province in year t ($\bar{p}_{-jmt}, \bar{x}_{-jmt}$). In addition, $\mathbf{Z}_{jmt}$ also contains interactions between each of these excluded instruments with the MSO dummy, as well as the four large firm dummies in $LargeFirm_{mt}$.

Defining $g(\theta) = \frac{1}{N_{jmt}} \sum_{j,m,t} \mathbf{Z}'_{jmt} \xi_{jmt}(\theta)$ as the empirical analogue to the population moment conditions, the structural parameter estimates are the solution to the following minimization problem:

$$\hat{\theta} = \arg\min_{\theta} \; g(\theta)' \Lambda g(\theta)$$

where Λ is a symmetric weighting matrix. I use a 2-step GMM estimation procedure to obtain efficient estimates where in the first step I obtain a consistent estimate of θ with $\Lambda = (\sum_{j,m,t} \mathbf{Z}'_{jmt} \mathbf{Z}_{jmt})^{-1}$. Using the first step estimate, I construct an optimal weighting matrix for Λ, and find the efficient second step GMM estimate of θ. See online Appendix B for computational details.

3.3. *Identification.* The model is identified from the assumed population moment conditions. Importantly, they assume that the excluded instruments in $\mathbf{Z}_{jmt}$ are uncorrelated with the product-, market- and year-specific unobserved demand shock for non-content product quality, ξ_{jmt}. Given this assumption, exogenous variation in prices and channel counts induced by the excluded demand instruments can be used to identify α, β_1, and β_2. This instrumenting strategy attempts to correct for simultaneity bias in estimating these parameters; this naturally arises if firms account for unobserved demand shocks in making their pricing and bundling decisions.

For identification, I use three sources of exogenous variation in prices and channel counts. The first arises from variation in cable system characteristics, including its size, channel capacity, kilometers of

[29]The CRTC Master Files report estimates of each cable system's potential number of subscribers, which includes households currently connected to the cable system (i.e., the number of homes passed) and those who are not connected to the cable system but who could potentially be connected. Channel capacity corresponds to the total number of basic and non-basic channel signals that can be transmitted from the cable system, and kilometers of cable is the total amount of cable needed to connect all the homes that are currently passed by the cable system. Collectively, these three variables characterize the current size and infrastructure inherent in a license's cable system.

cable, and the firm size of a cable system's current operator. These variables exhibit large within-license variation as a result of the consolidation process. As was discussed in Section 2, cable mergers often saw larger firms acquire smaller firms' cable systems, leading to upgrades and expansion of these systems' infrastructure together with alterations to basic and non-basic cable prices and bundles. This variation can be exploited for identification under the assumption that, controlling for firm/province/year fixed effects and local demographics, changes in cable prices and channel counts related to changes in firm size and cable system infrastructure are independent of local demand shocks, ξ_{jmt}.

The instrumenting strategy also exploits variation in prices and channels counts due to the different regulatory classes across cable systems. Recall that these classes are strictly defined by a market's population, which is presumably exogenous to unobserved demand shocks for non-content cable quality. Thus, the $Class2_{mt}$ and $Part3_{mt}$ variables are valid instruments under the assumption that basic cable price caps do not directly affect consumers' choices, but rather have an indirect effect through their impact on cable companies' pricing and bundling decisions.

The final two instruments for $\bar{p}_{-jmt}$ and $\bar{x}_{-jmt}$ are average prices and channel counts for bundle j across all *other* markets in market m's province in year t. The are often called "Hausman" instruments (Hausman **?**). They have been used in previous studies where market-level panel data are available, such as Nevo (2001) or Crawford's and Yurukoglu's (2012) recent study of the welfare impact of àla carte bundling in cable TV. If we assume ξ_{jmt} is uncorrelated across markets, then prices and channel counts from other markets are valid instruments for p_{jmt} and x_{jmt}. Given firm fixed effects are controlled for, transitory firm-level cost shocks can generate correlation in cable prices and channel counts across markets and over time. Thus, these shocks can be exploited for identification under this assumption.

There are at least two caveats with this identification strategy. First, if cable companies target licenses with large unobserved demand shocks for acquisitions, then instruments based on cable system characteristics and firm size would be invalid. I suspect this is not a major issue for mergers among MSOs, which affect many licenses and generate a large portion of the exogenous variation in cable prices and channel counts. These mergers are negotiated at a national level, and while they may depend on regional demand or macroeconomic shocks (which the province and year fixed effects control for), it is unlikely they depend on local unobserved non-content cable quality. In this sense, the identification strategy is similar to that from Hastings (2004) or Houde (2012) who identify the causal impact of competition/market structure on equilibrium prices by assuming mergers are negotiated nationally and do not depend on local unobserved demand shocks.

It is plausible that mergers between small cable companies in contiguous licenses may depend on local demand shocks, thereby undermining the identification assumption. However, these mergers tend to yield little identifying variation in terms of changes in firm size or cable system infrastructure (relative to the many mergers involving national cable companies like Rogers and Shaw). As such, mergers between small companies likely have a small impact on the parameter estimates.

The second issue is whether unobserved non-content cable quality is spatially correlated across markets; indeed, this is a potential problem for any study that uses Hausman instruments to identify demand. Here I rely on the same arguments as Crawford and Yurukoglu (2012): like their U.S. counterparts, cable markets in Canada during this period were physically distinct locations, and local cable companies had a great deal of authority in managing local service quality (which is likely the main driver of the unobserved ξ_{jmt} shocks).

To see how the instruments identify the demand parameters, consider the identification α. Higher values of α imply consumers are more sensitive to exogenous changes in basic and non-basic cable prices. Thus, I can identify α by observing how market shares for basic and non-basic services in the data change in response to exogenous variation in prices generated by the excluded demand instruments.

The firm-specific channel quality coefficients for basic and non-basic cable, β_1 and β_2, are a key novelty of the demand specification. As discussed, these coefficients capture consumers' relative tastes for basic and non-basic channels across the many licenses owned by a given national cable company. For example, if β_1 increases relative to β_2 for Rogers, then consumers in Rogers's licenses increasingly prefer additional basic channels to non-basic channels. Thus, the observed rate at which Rogers's licenses' market shares for basic and non-basic cable change in response to exogenous variation in basic and/or non-basic channel counts (induced by the demand instruments) identifies the relative magnitudes of β_1 and β_2 for Rogers. Similar logic applies for the identification of β_1 and β_2 for non-major MSOs, Shaw, Cogeco, and Vidéotron.

Differences in the rate at which market shares change in response to exogenous changes in channel counts *across* licenses operated by different large cable companies, or by MSO and non-MSO operators, also helps identify the coefficients in β_1 and β_2. Keeping with the example, if Rogers offers higher unobserved per-channel programming quality in their bundles than other companies (single license operators, non-major MSOs, Shaw, Cogeco, Vidéotron), then, all else equal, we should see larger increases in Rogers's basic or non-basic market shares in response to equivalent exogenous increases in basic or non-basic channel counts that we see for other companies.

The remaining demand parameters, ψ_0, ψ_1, ψ_2, are identified from variation in exogenous demand covariates (demographics, dummies for firms, years, and provinces) and variation in cable prices, bundle qualities, and market shares. For example, suppose two licenses, A and B, have the same basic and non-basic cable prices and bundle qualities, but A has higher urban density, lower non-basic shares, and higher basic shares. This variation would identify a negative ψ_1 coefficient for urban density, or equivalently, a lower scale distribution of unobserved cable quality taste shocks for more urban licenses. In words, this variation would identify a negative effect of urban density on cable demand.

3.4. *Results.* The demand parameter estimates are reported in Table 3. The estimated price coefficient of $\hat{\alpha} = -0.524$ compares to an OLS estimate of -0.314 (std. error = 0.164), which highlights the importance of accounting for the endogeneity of prices to unobserved non-content cable quality demand shocks.

[TABLE 3 ABOUT HERE]

Using the estimated model, I can compute price elasticities of demand for basic and non-basic cable. These figures are often of independent interest, particularly in Canada where, at least to my knowledge, there are no previously published elasticity estimates. In markets where both basic and non-basic cable are offered, the median own price elasticities of demand for basic and non-basic cable are -4.21 (s.d.=3.09) and -5.45 (s.d.=7.56). In contrast, rural markets with only basic cable have smaller price elasticities with a median of -1.30 (s.d.=2.15). As suggested above, the difference in the basic demand elasticities across these market-types is potentially due to the fact that consumers in larger urban markets with both basic and non-basic cable have relatively more leisure options (such as movie theatres or live sports) and hence more elastic demand for cable than consumers in rural markets with only basic cable.

There are a number of studies that estimate cable demand elasticities in the U.S. that serve as useful benchmarks. The closest relevant analysis in terms of the demand model and data is Crawford and Yurukoglu (2012).[30] They report median own price elasticities of -4.12, -6.34, and -13.11 for basic, expanded basic, and digital basic in markets where these three cable packages are offered. Perhaps not surprisingly, our results together suggest remarkably similar basic demand elasticities between Canadians and Americans. As Crawford and Yurukoglu (2012) note, the magnitude of these estimates are well within the range of demand elasticity estimates from numerous other studies. These include basic price elasticity estimates of

[30] I cannot directly contrast my estimated elasticities to those from Chu (2010) since he does not report them.

-1.5 (Rubinovitz 1993), -2.19 (FCC 2002), and -5.9 (Chipty 2001).[31] Unfortunately, I cannot find a directly comparable numbers for rural markets with only basic cable; however, the -1.30 median elasticity for these markets also falls within the range of elasticity estimates from these studies.

The cable quality coefficients estimates in the table highlight two main results: (1) non-basic channels are perceived to have higher quality than basic channels; (2) the dominant MSOs offer higher per-channel basic and non-basic channel quality than smaller cable companies, with the quality differential being more pronounced for non-basic cable.

These patterns are highlighted in Table 4, where I report sample averages (across license-years) of the mean and standard deviation of willingness to pay (WTP) for basic and non-basic channels in licenses with single-system operators, non-major MSOs, and the four dominant MSOs.[32] For instance, the mean WTP for consumers in rural licenses with single-system operators is \$0.48 and \$0.50 (per month) for an additional basic and non-basic channel. In contrast, consumers in Shaw's licenses have a mean WTP of \$0.61 and \$1.08. Notice, however, that among the dominant MSOs there are some differences in the relative quality of basic and non-basic channels: consumers in Rogers's licenses tend to perceive per-channel basic cable quality as being quite low, while Vidéotron's consumers in Québec perceive basic channels as having relatively higher per-channel quality than non-basic channels.

[TABLE 4 ABOUT HERE]

The remaining taste distribution parameters suggest important heterogeneity in consumers' preferences for cable quality. This can be seen in Table 4, where the sample average of standard deviation in consumers' WTP for basic channels ranges from \$0.03 to \$0.12, and from \$0.06 to \$0.14 for non-basic channels. In general the estimates yield more disperse preference for cable quality in larger, more urban cable markets.

[31] Similarly, Goolsbee and Petrin (2004) report an elasticity of -1.5 for extended basic cable, and also find demand is more elastic for higher quality cable packages like premium cable or Direct Broadcast Satellite.

[32] Consumer i's WTP for an extra channel in bundle j for MSO k can be computed as $\text{wtp}_{ijkmt} = \frac{t_i \times (z_{mt}\beta_{jk})}{\alpha}$. By simulating wtp_{ijkmt} values I can obtain estimates of the mean and standard deviation of WTP for a given MSO for each license-year in the sample. The WTP figures reported in Table 4 are sample averages from the model's predicted WTP distributions in each market. They do not account for estimation error in the parameter estimates in computing market-level WTP means and standard deviations.

4. SUPPLY

With an estimated demand model in hand, I now develop the supply-side of the model. It also falls within the differentiated products framework of Berry, Levinsohn, and Pakes (1995). The model entails two key assumptions regarding firms' costs and conduct: they have constant marginal costs and maximize short-run profits within their local markets, subject to regulatory constraints. This framework has been previously applied to the cable TV industry (Chu 2010; Crawford and Yurukolu 2012), as well as many other industries such as automobiles, banking, retail gasoline, and personal computers.

Having developed the supply-side model, I then use my unique cost data to test the validity of the model's assumptions regarding costs and conduct. The test involves three steps:

1. Recover the marginal costs for basic and non-basic cable implied by the supply-side model for each licence-year, given the estimated demand system, and observed prices, channel counts and product market shares. These inferred costs are what researchers typically use when cost data are unavailable.

2. Directly estimate marginal costs using the cost data. Two sets of estimates are used: (1) average variable costs, which correspond to marginal costs under the constant marginal cost assumption (which is common in the empirical literature on differentiated products markets); and (2) estimates based on a flexible translog cost function that does not assume constant marginal costs.

3. Compare the distributions of the inferred and directly estimated marginal costs from **1.** and **2.** Such a comparison is a valid test of the supply-side modeling assumptions since the demand model was estimated without imposing any supply-side restrictions.

I now develop the supply-side model, describe the marginal cost estimation strategies from **1.** and **2.**, and present the results from **3.**

4.1. *Model.* The model assumes a constant per-subscriber marginal cost c_{jmt} of providing cable bundle j to an additional household in license m in year t. These costs are specified as follows:

$$\begin{aligned} c_{1mt}(\mathbf{w}_{mt};\gamma) &= x_{1mt} \cdot c^x_{1mt} + c^{nx}_{1mt} \\ &= x_{1mt} \cdot (w_{1mt}\gamma^x_1 + \omega^x_{1mt}) + w_{1mt}\gamma^{nx}_1 + \omega^{nx}_{1mt} \end{aligned} \tag{6a}$$

$$\begin{aligned} c_{2mt}(\mathbf{w}_{mt};\gamma) &= c_{1mt}(\mathbf{w}_{mt};\gamma) + \Delta c_{2mt}(\mathbf{w}_{mt};\gamma) \\ &= c_{1mt}(\mathbf{w}_{mt};\gamma) + x_{2mt} \cdot c^x_{2mt} + c^{nx}_{2mt} \\ &= c_{1mt}(\mathbf{w}_{mt};\gamma) + x_{2mt} \cdot (w_{2mt}\gamma^x_2 + \omega^x_{2mt}) + w_{jmt}\gamma^{nx}_2 + \omega^{nx}_{2mt} \end{aligned} \tag{6b}$$

where c^x_{jmt} is a per-channel, per-subscriber programming content cost, and c^{nx}_{jmt} is a per-subscriber non-content cost relate to labor and administrative expenses. These marginal cost components depend on content and non-content cost parameters (γ^x_j and γ^{nx}_j), a set of common content and non-content cost shifters (w_{jmt}), and cost shocks (ω^x_{jmt} and ω^{nx}_{jmt}). The non-basic marginal cost specification in (6b) reflects the fact that non-basic purchases are tied to basic purchases. Thus, $\Delta c_{2mt}(\mathbf{w}_{mt};\gamma)$ represents the incremental portion of the marginal cost of providing non-basic cable, over and above the basic marginal cost. I collect all cost parameters with the vector $\gamma = \{\gamma^x_1, \gamma^x_2, \gamma^{nx}_1, \gamma^{nx}_2\}$.

The constant content and non-content marginal costs assumptions are pervasive among empirical studies of the cable TV industry. Various authors have justified these assumptions based on independent policy studies of the cable TV industry.[33] Content costs are mainly driven by affiliation payments, which cable companies and channel providers negotiate over. These payments mainly depend on a per-subscriber affiliate fee which a cable company pays to a channel provider for including the provider's channel in one of its cable packages. Hence, the total per-subscriber content cost c^x_{jmt}, which largely correspond to these upstream payments to channel providers, is likely to be constant across subscribers.

Total non-content variables costs are related to labor and administrative expenses, and are similarly as-

[33]For instance, Rubinovitz's (1993) seminal study cites Smiley's (1986) report on the cable industry for the U.S. Department of Justice, which finds that a cable system's variable costs are proportional to the number of subscribers. Other studies that argue for the constant cable TV marginal cost assumption include Otsuka and Mayo (1991), Chu (2010), and Crawford and Yurukoglu (2012).

sumed to scale with the number of subscribers in a market; indeed, this appears to be the case with my cost data. Thus, it is reasonable *a priori* to assume that c^{nx}_{jmt} is constant across subscribers as well.

My choice of variables to include in w_{jmt} (for $j = 1,2$) is motivated by the empirical patterns and discussion from Section 2. Importantly, they include $FirmSize_{mt}$, which is the size of the cable company in license m and year t in terms of the number of subscribers it served nationally in year $t-1$. In light of the discussion and descriptive statistics from Section 2.3, I expect larger cable companies to negotiate lower per-channel per-subscriber content costs. I also include the MSO dummy MSO_{mt} and the vector of individual MSO fixed effects $Firm_{mt}$ in w_{jmt}. These will account for cost-reducing vertical integration effects among MSOs and some channel providers. These fixed effects will also control for any firm-specific differences in content costs (beyond firm size effects), possibly due to differences in cable bundle composition across firms.[34] I also include characteristics of the cable system in w_{jmt} that may affect the marginal cost of providing content to subscribers locally. These include urban density (Urb_{mt}), average household income (Inc_{mt}), total potential system size ($SysSize_{mt}$), and channel capacity ($ChanCap_{mt}$). For any regressions based on basic costs, w_{jmt} also includes a dummy variable that equals one if a license only has basic cable. Finally, I include year and province dummies in w_{jmt} to account for any unobserved differences in content costs across years or provinces.

I similarly expect firm size and MSO fixed effects to be important drivers of firms' non-content costs. Both variables will capture any firm-specific non-content cost efficiencies (or inefficiencies). For example, if dominant cable companies are more experienced and effective in managing labor and operating costs within their markets, then their fixed effects should have negative γ^{nx}_j coefficients.

Like the β_1 and β_2 parameters in the cable quality functions, γ^x_j is also a reduced-form parameter vector. Fundamentally it is a function of two interrelated supply-side processes: (1) cable companies' optimal bundling decisions; and (2) bilateral negotiations between cable companies and upstream channel providers over affiliation payments. Again, my lack of channel identity data prevents me from explicitly modeling these features of the industry. Further, the complexity of cable companies' optimal bundling problems would only be exacerbated if bilateral negotiations were jointly modeled. Thus, even if channel identity data were available, it is not clear whether it would be computationally feasible to conduct merger simulations in Section 5 without substantial simplifications to, or approximations of, these joint bundling and bargaining

[34]For example, if a dominant cable company like Rogers offers higher quality non-basic channels on average that demand higher affiliation payments, then, all else equal, we might expect the Rogers dummy coefficient in γ^x_2 to be positive.

problems.

Regardless, my reduced-form approach to modeling the impact of firm-specific differences on content costs introduces limitations. Because they are reduced-form, all of the parameters in γ_j^x are only valid for studying the industry during the 1990-1996 period. They cannot be used to make any out-of-sample predictions over longer-time horizons, such as the impact channel entry/exit would have on cable bundle quality and negotiated content costs.

Moreover, the reduced-form specification complicates the interpretation of differences in bundle quality and content costs across firms. Per-channel content quality for a given firm is revealed by the estimated β_1 and β_2 quality coefficients from the demand model. If firms that offer higher per-channel quality pay higher per-channel content costs, then these cost differences will be reflected in the estimated firm fixed effects in γ_1^x and γ_2^x. Conditional on these cost effects, the impact of $FirmSize_{mt}$ on content costs in equations (6a) and (6b) largely determines the number of channels a firm includes in its cable packages. So, while the model accounts for firm-specific effects of program-content quality on demand and content costs, it requires an interpretation that is more involved than if channel identity data were available.

4.1.1. *Profit maximization and regulation.* The cable company in license m and year t knows the distributions of t_i and ε_{ijmt}, but not the vertical types or horizontal preference shocks of individuals. As a local monopolist, the company chooses prices and channel counts to maximize profits subject to the regulator's basic price and must-carry regulations, as well as the tying constraint:

$$\max_{\{(p_{jmt},x_{jmt})\}_{j=1}^{J_{mt}}} \quad \pi_{mt} = \sum_{j=1}^{J_{mt}} \Big[(p_{jmt} - c_{jmt}(\mathbf{w}_{mt};\gamma)) s_{jmt}(\mathbf{p}_{mt},\mathbf{x}_{mt};\theta) Q_{mt} \Big] \tag{7}$$
$$\text{subject to} \quad p_{1mt} \le \bar{p}_{1mt}; \quad x_{1mt} \ge \underline{x}_{1mt}; \quad p_{2mt} = p_{1mt} + \Delta p, \;\; \Delta p \ge 0$$

where $\bar{p}_{1mt}$ is the basic cable price cap, and $\underline{x}_{1mt}$ is the number of channels that must be included in the basic package under the must-carry regulations. Recall that price caps only apply to Class 1 licenses (i.e., markets with more than 6000 subscribers) and Class 2 licenses (i.e., markets with 2000 to 6000 subscribers).[35] The number of packages offered in license m at time t, J_{mt}, is taken as exogenous. While the number of channels vary over time within licenses (particularly following acquisitions), the number of packages rarely changes,

[35]The quality and marginal cost specifications in equations (2a),(2b),(6a),(6b) already incorporate the tying requirements for basic and non-basic channels in the demand and cost equations.

so I take J_{mt} as given.

For completeness, I have listed all the regulatory constraints that cable companies face within their licenses. However, as in Chu (2010), Crawford and Yurukoglu (2012), there is little reason to believe the channel carriage constraints will bind. From my reading of the CRTC Decision and Notices regarding must-carry channels, I suspect that most licenses have basic cable packages that exceed the minimum quality standard. While I do not have channel identity data to confirm this, a reasonable estimate of the minimum number of must-carry channels included in a basic package in the sample is six: CBC, CTV, Global, CPAC, an educational channel, and a local community channel.[36] Given that 99% of all license-years have 10 or more basic channels, it is unlikely these quality constraints bind in many markets.

In contrast, a number of studies have shown that basic price regulation can affect prices (Rubinovitz 1993; Mayo and Otsuka 1991; Kelly and Ying 2003). I account for and assess the impact of these constraints on the empirical results below.

Beyond the role of regulation, another specification issue is the degree to which pricing and channel bundling decisions are centrally determined for large MSOs like Rogers or Shaw. This goes against my assumption that companies determine prices and cable packages at the license-level. Such centralization in decision-making could introduce bias into the model's implied marginal costs and simultaneity bias into the cost function parameter estimates in equations (6a) and (6b).[37] This is particularly important when estimating marginal costs without cost data (as I discuss in a moment), where the license-level first order conditions are assumed to hold. While I do not test for centralization in pricing and bundling decisions,

[36]The first three channels correspond the country's three major national broadcast networks and their regional affiliates that have existed since well before the 1990s. CPAC was launched in 1992. Large urban markets will likely have more local channels offered under the must-carry provisions (such as CITY-TV in Toronto), however, it is extremely unlikely the minimum basic cable quality standard binds in these markets since they have more than 20 basic channels on average.

[37]A plausible scenario is if large MSOs set their prices and mark-ups based on a single estimate of the average price elasticity of demand across all of their licenses, and not license-level demand elasticities. Then, any licenses that have estimated a small price elasticity of demand for (such as rural licenses) will overstate the true mark-ups and understate the true marginal costs. Similarly, licenses with high elasticities will underestimate mark-ups and overestimate marginal costs. This would introduce simultaneity bias in my parameter estimates in equations (6a) and (6b). Keeping with the example, if true marginal costs are indeed larger in more rural markets, but mark-ups are determined centrally based on an aggregate demand elasticity estimate, then the magnitude of my estimate of the impact of urban density on marginal costs (through either its content or non-content component) will be biased upward as the specification error from assuming local and not centralized pricing will predict too large a (negative) correlation between urban density and marginal costs.

I find there is large variation in cable prices and channel counts across the cable licenses of larger MSOs, which is related to local demand and cost variables. This suggests that local market conditions do indeed play an important role in how cable companies determine their cable packages.

4.2. *Estimating marginal costs without cost data.* The conventional approach to estimating marginal costs when cost data is unavailable is to assume a specific form of conduct, and then recover the marginal costs from the first order conditions that govern firms' behavior. That is, one finds the marginals costs that perfectly rationalize firms' observed pricing decisions (or any other endogenous variables of interest), given the conduct assumption.

Assuming the cable company in license m and year t maximizes short-run profits by choosing prices and qualities for its J_{mt} cable packages, there are $2 \times J_{mt}$ first order equations that govern its behavior. These equations can derived by differentiating the Lagrangian that corresponds to the constrained optimization problem in (7) with respect to basic and non-basic prices and channel counts.[38] For Part 3 licenses (i.e., markets with less than 2000 subscribers) that do not face price regulation, the first order conditions with respect to prices and channel counts can be written in matrix form as:

$$\tilde{\mathbf{s}}_{mt} + \Delta_{\mathbf{p}_{mt}}(\mathbf{p}_{mt} - \mathbf{c}_{mt}) = \mathbf{0} \tag{8a}$$

$$-\tilde{\mathbf{s}}_{mt} * \mathbf{c}^{x}_{mt} + \Delta_{\mathbf{x}_{mt}}(\mathbf{p}_{mt} - \mathbf{c}_{mt}) = \mathbf{0} \tag{8b}$$

where as before all boldface elements are $J_{mt} \times 1$ vectors that stack the elements for each cable package in a market, with the exception of the $\tilde{\mathbf{s}}_{mt}$ vector which equals s_{1mt} in markets with basic cable only, and $[s_{1mt} + s_{2mt},\ s_{2mt}]'$ in markets with basic and non-basic cable. The '$*$' operator denotes the Hadamard product, and $\Delta_{\mathbf{p}_{mt}}$ is a $J_{mt} \times J_{mt}$ matrix whose jk^{th} entry is $\Delta_{jk} = \partial s_{jmt}/\partial p_{kmt}$ (and similarly for $\Delta_{\mathbf{x}_{mt}}$). These latter matrices of derivatives with respect to prices and channel counts are computed using the estimated demand system.

[38]Ignoring the channel carriage constraint and non-negativity constraint on Δp from (7) (i.e., assuming that non-basic prices differ from basic prices), the Lagrangian for this constraint optimization problem is given by:

$$\max_{\{(p_{jmt}, x_{jmt})\}_{j=1}^{J_{mt}}} \mathscr{L}_{mt} = \sum_{j=1}^{J_{mt}} \Big[(p_{jmt} - c_{jmt}(\mathbf{w}_{mt};\gamma))s_{jmt}(\mathbf{p}_{mt}, \mathbf{x}_{mt};\theta)Q_{mt}\Big] - \lambda_{mt}(\bar{p}_{1mt} - p_{1mt})$$

where λ_{mt} is the Lagrange multiplier for the basic price constraint.

Notice that the first set of equations in (8a) can be inverted to obtain the usual "marginal costs equal prices less mark-ups" formula that is commonly used for identifying marginal costs with differentiated products models:

$$\mathbf{c}_{mt} = \mathbf{p}_{mt} + \Delta_{\mathbf{p}_{mt}}^{-1} \tilde{\mathbf{s}}_{mt} \equiv \mathbf{r}_{mt} \tag{9}$$

where $\mathbf{r}_{mt}$ is the $J_{mt} \times 1$ vector of marginal revenues implied by the observed prices and estimated demand system. In the context of unregulated local cable monopolies, the implied marginal costs from the model are simply these basic and non-basic marginal revenues.

To further identify the content and non-content components of marginal costs, I exploit the second set of first order conditions in (8b) and substitute in the J_{mt} components of $\mathbf{c}_{mt}$ with $c_{1mt} = x_{1mt} \cdot c^x_{1mt} + c^{nx}_{1mt}$ and $c_{2mt} = c_{1mt} + x_{2mt} \cdot c^x_{2mt} + c^{nx}_{2mt}$. This yields $2 \times J_{mt}$ equations, which can be use to recover the $2 \times J_{mt}$ unknowns, c^x_{jmt} and c^{nx}_{jmt} for $j = 1 \ldots J_{mt}$. Using the inferred content and non-content costs, the marginal cost parameters for can be estimated using the following linear regressions:

$$c^x_{jmt} = w_{jmt} \gamma^x_j + \omega^x_{jmt} \tag{10a}$$

$$c^{nx}_{jmt} = w_{jmt} \gamma^{nx}_j + \omega^{nx}_{jmt} \tag{10b}$$

where the j subscripts on γ^x_j and γ^{nx}_j indicate that I estimate separate content and non-content cost functions for basic and non-basic cable.

4.2.1. *Accounting for price regulation.* The first order conditions in (8a) potentially do not hold in price-regulated Class 1 or Class 2 licenses. In these markets, the pricing first order conditions are:

$$\tilde{\mathbf{s}}_{mt} + \Delta_{\mathbf{p}_{mt}} (\mathbf{p}_{mt} - \mathbf{c}_{mt}) - \boldsymbol{\lambda}_{mt} = \mathbf{0} \tag{11}$$

where $\boldsymbol{\lambda}_{mt} = \lambda_{mt}$ if $J_{mt} = 1$, $\boldsymbol{\lambda}_{mt} = [\lambda_{mt}, 0]'$ if $J_{mt} = 2$, and λ_{mt} is the Lagrange multiplier on the price-cap constraint. Inverting these equations, we obtain: $\mathbf{r}_{mt} = \mathbf{c}_{mt} + \Delta_{\mathbf{p}_{mt}}^{-1} \boldsymbol{\lambda}_{mt}$. From this expression we see that if one assumes firms price according to the first order equations in (8a), but in reality basic price caps are binding in some markets, then estimates of $\mathbf{c}_{mt}$ based on $\mathbf{r}_{mt}$ alone will be biased downwards. Because non-basic and basic cable purchases are tied, the implied (total) marginal costs for both basic and non-basic cable will exhibit such a bias.

Following a similar strategy to Goldberg (1995), I can recover an estimate of basic and non-basic marginal costs in the presence of regulatory constraints by exploiting cross-license variation in regulatory status. Specifically, I use the following regression to purge the marginal revenue estimates of regulatory effects:

$$r_{jmt} = 1\{Class^1_{mt}\}\delta_{1j} + 1\{Class^2_{mt}\}\delta_{2j} + w_{jmt}\gamma_j + \omega_{jmt} \tag{12}$$

where r_{jmt} is the j^{th} element of $\mathbf{r}_{mt}$, and $1\{Class^k_{mt}\}$ is an indicator variable that equals one if license m in year t is subject to Class k basic price regulation. Assuming all other market-level variables that affect marginal costs are controlled for, the OLS estimate of δ_{kj} will correspond to the average value of the j^{th} element of $\Delta^{-1}_{\mathbf{p}_{mt}}\lambda_{mt}$ across license-years with Class k price regulation. An estimate of the marginal cost can then be recovered as the residual value of r_{jmt} after purging it of regulatory effects: $\hat{c}_{jmt} = w_{jmt}\hat{\gamma}_j + \hat{\omega}_{jmt}$, where $\hat{\gamma}_j$ and $\hat{\omega}_{jmt}$ is the OLS estimate and residual from (12).[39]

4.2.2. *Supply-side residuals.* The above discussion highlights a more general problem in inferring firms' marginal costs using differentiated products models. Given the estimated demand system, the inferred $\mathbf{c}_{mt}$ vector from (9) only corresponds to marginal costs if the conduct assumption of profit maximization is correct. If the impact of price regulation on firms' pricing decisions was ignored, the marginal cost estimates, and any related policy analyses, would be biased. Even if I account for the impact of price regulation using the regression in (12), there is no guarantee that conduct misspecification will not undermine the interpretation of $\mathbf{c}_{mt}$ as marginal costs.

These issues are not unique to the current setting; indeed, the pricing equations in (9) are almost identical to the first order equations from differentiated Bertrand oligopoly models, which are typically used for recovering marginal costs. Any sources of conduct misspecification, such as firms not strictly pricing according to a Nash Equilibrium, would undermine the interpretation of $\mathbf{c}_{mt}$ as marginal costs. Borrowing from Peters (2006), $\mathbf{r}_{mt}$ can generally be referred to as a *supply-side residual*: it is the residual difference between observed prices and predicted mark-ups from the demand model. If the researcher's conduct assumption is correct, the residual will correspond to firms' marginal costs, as per equation (12). This residual-based

[39]See section 5.5 of Goldberg (1995) for further discussion about identifying marginal costs in a differentiated products model in the presence of regulatory constraints. Given the $\hat{c}_{jmt}$ estimate, as before, we can identify the content and non-content components of marginal costs using the (unrestricted) channel count first order conditions.

interpretation of the model's implied costs proves useful below, when I compare the marginal costs implied by the model to those estimated using cost data.

4.3. *Estimating marginal costs with cost data.* I can directly estimate firms' marginal costs with the unique cost data. These are necessarily estimates since the CRTC collects and audits cable companies' total costs and not their marginal costs. For the purpose of validating the supply-side model, I take two approaches to estimating these costs.

4.3.1. *Average variable costs.* The first approach is based on the common constant marginal costs assumption. Under this assumption marginal costs can be approximated by average monthly variable costs:

(13a) $$\hat{c}_{1mt} = (C^{x}_{1mt} + C^{nx}_{1mt}) \Big/ (12 \cdot \sum_{j=1}^{J_{mt}} Q_{jmt})$$

(13b) $$\hat{c}_{2mt} = \hat{c}_{1mt} + (C^{x}_{2mt} + C^{nx}_{2mt}) \Big/ (12 \cdot Q_{2mt}) = \hat{c}_{1mt} + \Delta\hat{c}_{2mt}$$

where C^{x}_{jmt} and C^{nx}_{jmt} are total annual content and non-content costs[40] for package j, Q_{jmt} is the number of households that subscribe to package j, and $\Delta\hat{c}_{2mt}$ denotes the incremental part the marginal cost for non-basic cable above beyond basic marginal cost.[41] Throughout I use the convention that marginal costs estimates based on cost data have a " ˆ ", while marginal cost estimates inferred from the supply-side first order conditions do not.

It is also straightforward to compute the empirical analogues to the content and non-content components

[40]Recall from Section 2 and Table 1 that total content costs are reported as technical expenses plus affiliation payments, while total non-content costs are the sum of labor and sales and administrative expenses.

[41]These calculations reflect the fact that the CRTC Master Files report total cost components for all subscribers to a particular cable tier in a market. So, for example, in markets where basic and non-basic cable are offered, C^{x}_{1mt} is the total amount of basic content costs paid for serving the Q_{1mt} consumers who only subscribe to basic cable *and* for the Q_{2mt} consumers who subscribe to non-basic cable (who must subscribe to basic cable because of the tying constraint). The Master Files further report C^{x}_{2mt}, which is the total *incremental* non-basic content cost, which is paid to provide non-basic services to the Q_{2mt} non-basic subscribers. The reporting is similar for total non-content basic and non-basic costs, C^{nx}_{1mt} and C^{nx}_{2mt}.

of marginal costs for basic and non-basic cable under the constant marginal cost assumption:

$$\hat{c}^{x}_{1mt} = C^{x}_{1mt} \Big/ (12 \cdot x_{1mt} \cdot \sum_{j=1}^{J_{mt}} Q_{jmt}) \tag{14a}$$

$$\hat{c}^{x}_{2mt} = C^{x}_{2mt} \Big/ (12 \cdot x_{2mt} \cdot Q_{2mt}) \tag{14b}$$

$$\hat{c}^{nx}_{1mt} = C^{nx}_{1mt} \Big/ (12 \cdot \sum_{j=1}^{J_{mt}} Q_{jmt}) \tag{14c}$$

$$\hat{c}^{nx}_{2mt} = C^{nx}_{2mt} \Big/ (12 \cdot Q_{2mt}) \tag{14d}$$

4.3.2. *Translog cost function.* The second approach to estimating marginal costs relaxes the assumption of constant marginal costs. Specifically, I estimate a multi-product translog cost function (Christensen, Jorgenson, and Lau 1973), and use it to estimate firms' marginal costs. This function provides a second-order second-order approximation to an arbitrary cost function, and can be written as follows:

$$\begin{aligned}
\ln C_{mt} = {} & \vartheta_0 \ln X_{mt} + \sum_g \ln \vartheta_g W_{g,mt} + \sum_i \rho_i \ln Y_{i,mt} + \sum_k \varphi_k A_{k,mt} \\
& + 0.5 \sum_{gh} \vartheta_{gh} \ln W_{g,mt} \ln W_{h,mt} + 0.5 \sum_{ij} \rho_{ij} \ln Y_{i,mt} \ln Y_{j,mt} + 0.5 \sum_{k\ell} \varphi_{k\ell} \ln A_{k,mt} \ln A_{\ell,mt} \\
& + \sum_{gi} \tau_{gi} \ln W_{g,mt} \ln Y_{i,mt} + \sum_{gk} \tau_{gk} \ln W_{g,mt} \ln A_{k,mt} + \sum_{ik} \tau_{ik} \ln Y_{i,mt} A_{k,mt} + u_{mt}
\end{aligned} \tag{15}$$

where C_{mt} is the total cost of operating the cable system in license m and year t, W_{mt} and Y_{mt} are vectors of input prices and outputs, A_{mt} is a vector of cable system characteristics, X_{mt} is a vector of additional total cost shifters, $\vartheta, \rho, \varphi, \tau$ are parameters, and u_{mt} is an econometric error. Note that there is an abuse of notation: g, h, i, j, k, ℓ simply index the variables in each vector and are not to be confused with indexes for individuals, households or products or any other indices used elsewhere in the paper.

For the sake of brevity, and because the methods for estimating translog cost functions are well-established, I briefly discuss the variables included in the cost function here. Detailed discussions of variable definitions, estimation, and parameter estimates is provided in online Appendix B.[42] The variables in W_{mt} include input prices for labor, programming content, and non-content "materials" (which correspond to administrative and sales costs). The outputs in Y_{mt} are the total number of subscribers for basic and non-basic cable (Q_{1mt} and

[42]The specification choices reflect those of Kelly and Ying (2003), who estimate translog cost function using proprietary cost data for U.S. cable systems from the Federal Trade Commission.

Q_{2mt}). The system characteristics in A_{mt} include the number of basic and non-basic channels (x_{1mt} and x_{2mt}), size of the cable company ($FirmSize_{mt}$), a dummy variable that equals one if the local cable company is an MSO (MSO_{mt}), urban density (Urb_{mt}), and average household income (Inc_{mt}). The total cost shifters in X_{mt} include province and year dummies.

When estimating marginal costs with a translog cost function one must decide whether to treat capital as a fixed or variable input. Essentially this boils down to whether one thinks a year corresponds to the short- or long-run for a cable company.[43] If we assume equation (15) corresponds to a long-run cost function, then all inputs are variable, and capital input prices and costs must be included in W_{mt} and in calculating total costs C_{mt}. In contrast, if capital is treated as fixed, then capital input prices and costs are not included. Rather than assuming a short- or long-run cost function, I simply report both short and long-run marginal cost estimates below. If cable companies treat capital as fixed when making their pricing and bundling decisions, then the marginal costs implied by the model will be closer to the short-run marginal cost estimates.[44]

With an estimated multi-product cost function in hand, I can directly compute the marginal cost for bundle j in license m and year t as:

$$\hat{c}_{jmt} = (\partial \hat{C}_{mt} / \partial Q_{jmt})/12 \tag{16}$$

where $\hat{C}_{mt}$ corresponds to the cost function in equation (15) evaluated at the parameter estimates. I divide the marginal cost estimates by 12 because the translog cost function is based on annual cost data, while the marginal costs obtained from the supply-side model are monthly.

[43] Slade (2009) also raises this issue, but does not have cost data to investigate it. To my knowledge, there is no standard treatment of time horizons in the literature, as no papers have attempted to combine traditional approaches to cost function estimation with a structural differentiated products model.

[44] A further difficulty in comparing the marginal costs implied by the structural model to long-run marginal cost estimates that are based on capital assets is that while the prior unambiguously corresponds to economic costs, the latter likely reflect accounting rates of return. Indeed, a secondary motive for the *New Empirical Industrial Organization* was to use economic models to identify marginal costs and market power rather than rely on firms' cost data that includes capital expenditures that reflect accounting and not economic costs (Bresnahan 1989). There is a large literature on this subject, with Fisher and McGowan (1983) being a key reference. This issue is less problematic (if at all) when estimating short-run marginal costs based on non-capital expenses that do not suffer from being subject to arbitrary accounting rules and rate of return calculations. In particular, costs reported for labor, sales and administration, and programming content are simply the amount spent on these items year-to-year.

4.4. *Validating the supply-side model.* Table 5 presents sample medians for the implied marginal costs estimated from the supply-side model, and for the marginal cost estimates based on the cost data.[45] 95% percentile bootstrap confidence intervals are also reported. Importantly, these intervals, as well as all other confidence intervals and standard errors reported in the remainder of the paper, account for within-market dependence in marginal costs, and variability in inferred costs associated with the first-step demand parameter estimates. online Appendix C explicitly describes all of the bootstrapping procedures used in constructing the various confidence intervals and standard errors reported in Tables 5 through 10.

The validation results from Table 5 demonstrate that the estimated (total) marginal costs based on the data (Panel A) differ considerably from those based on the supply-side model (Panel B). The first set of marginal costs in Panel B correspond to the naive estimates from inverting equation (8a) for all products, markets, and years (i.e., those that correspond to the standard approach to identifying marginal costs). These marginal costs exhibit the largest bias. This should not be surprising is the context of a price-regulated industry. However, two addition findings suggest that, beyond the impact of regulation, the conduct assumption of profit maximization gives rise to specification error.

[TABLE 5 ABOUT HERE]

First, the estimates in the third row of Panel B, which account for the impact of basic price regulation, are still quite small compared to the estimates in Panel A. After controlling for regulatory effects on supply-side residuals using the regression in (12), the structural model yields marginal costs with a downward bias. This suggests that a number of firms are not profit-maximizing and could increase their profits by raising their prices. Nevertheless, these cost estimates are an improvement over the naive estimates that ignore basic price cap effects. Indeed, the difference in the two sets of supply-side residuals in Panel B is evidence that basic price caps are binding in some markets.[46]

Second, the differences in the marginal cost estimates are the largest in *unregulated* Part 3 licenses. This can be seen in the far right column of Table 5. The median supply-side residual/marginal revenue of 0.05

[45] I focus on medians to avoid the impact of extreme observations in comparing these different marginal cost distributions. The analysis is largely unchanged if I instead focus on sample means.

[46] The (unreported) δ_j^k coefficients for Class $k = 1$ and $k = 2$ licenses from the regression (12) for basic cable are $\hat{\delta}_1^1$=-8.53 (s.e.=2.34) and $\hat{\delta}_1^2$=-5.19 (s.e.=1.66) (standard errors are clustered at the license-level). The estimates for non-basic cable are $\hat{\delta}_2^1 = -3.52$ (s.e.=2.79) and $\hat{\delta}_2^2 = -0.76$ (s.e.=1.96).

for these licenses suggest that rural cable operators tend to be pricing on the inelastic portion of the demand curve. The bottom panel of the table highlights this fact: of the 613 unregulated licenses, 308 exhibit *negative* inferred marginal costs, or equivalently, marginal revenues. These cable companies could increase their revenues and profits by raising their basic cable prices, and they would be completely unconstrained in doing so.

Figure 2 sheds further light on the degree to which these unregulated cable companies deviate from profit maximizing behavior. The figure plots, as a function of the utility function's price coefficient α, the fraction of these markets where the model and cost data would predict that: (1) marginal revenue/marginal cost prediction is negative; (2) marginal revenue is less than marginal cost; and (3) marginal revenue is within \$5.00 of marginal cost.[47] The figure indicates that at the estimated price coefficient of $\hat{\alpha} = 0.52$, marginal revenue is less than marginal cost in 90% of the Part 3 licenses. The cable operators in these markets could increase profits by raising their prices. Even if α on extreme value of $\alpha = 2$ (implying a price elasticity of -8, which is two to four times typical estimates in the literature), we would still conclude that 70% of these firms are setting prices below profit-maximizing levels.

[FIGURE 2 ABOUT HERE]

The figure also shows that the estimated differences between marginal revenue and marginal cost are generally quite large. Marginal revenue is within a modest \$5.00 of marginal cost in only 20% of the unregulated licenses at the estimated α value. These optimization errors ultimately undermine and bias the marginal cost estimates that are implied by the supply-side model.

I also contrast the content and non-content marginal cost components implied by the model and cost data. Table 6 presents their corresponding summary statistics. The key finding is that the non-content component of marginal costs exhibits large biases, while the content component does not. In fact, the per-subscriber content costs implied by the supply-side model are quite close to those from the data for markets with only basic cable. These results suggest that firms' sub-optimal pricing decisions, rather than their channel count (or bundle size) decisions, drive the magnitude of the biases in the model's inferred marginal costs.

[TABLE 6 ABOUT HERE]

[47] I use the average variable cost estimate of marginal costs to construct this figure. The results are similar if I use the marginal cost estimates based on the translog cost function.

4.4.1. *Formally testing for differences in marginal cost estimates.* The results from Tables 5 and 6 indicate that the marginal cost estimates based on the cost data and the structural model differ considerably. However, such a simple comparison of summary statistics is not a formal test whether differences exist in a statistical sense. I undertake such a test by running (bootstrap) paired two-sample t-tests for differences in means for all combinations of data- and structural-model based marginal cost estimates from panels A and B in Tables 5 and 6. Tables A.1 and A.2 in online Appendix A report the results from these tests for each of the four sub-samples from Tables 5 and 6, and for each year of the sample.[48] The results confirm that the differences in marginal cost estimates based on the cost data and the structural model are indeed statistically significant.

4.4.2. *Caveats.* There are two important points to keep in mind when interpreting the finding of systematic underpricing of cable. First, this result is not due to complementary product pricing for other services like phone or Internet. Recall from the industry overview in Section 2 that cable companies do not yet offer these or any other services during the 1990-1996 sample period. Firms in the sample generate through cable services only.

Second, mis-specification of the demand system and/or demand instruments is a caveat to the validation results. This could also generate differences between the model's predicted marginal costs and those estimated using cost data for reasons unrelated to conduct mis-specification. For example, if I am underestimating true demand-side price sensitivity, then the supply-side model will overpredict mark-ups, and hence underestimate marginal costs.

While demand mis-specification is an important caveat, two results from above provide some reassurance that it does not undermine the validation results. First, recall that the estimated price elasticities of demand closely align with many prior elasticity estimates for the U.S. cable industry, which is very similar to the Canadian cable industry on both the demand and supply side. This includes estimates from Crawford and Yurukoglu (2012) who use a similar demand model, instrumenting strategy, and dataset. Second, recall

[48]The table presents the sample averages of the differences between the inferred marginal costs from the model and estimated marginal costs from the cost data. Bootstrap confidence intervals for these averages are also reported, as well as the results from testing whether the average differences are significantly different from zero (e.g., the paired two-sample t-test). I conduct the tests by sample year to ensure that the sample observations are i.i.d. This assumption is quite plausible across markets within a year. Analogous tests based on a pooled dataset would be undermined by within-license dependence in the marginal cost estimates over time.

from Figure 2 that the majority of unregulated markets would be predicted to have excessively low cable prices even at extremely high price elasticities of demand. Thus, even if my demand model under predicts price sensitivity to some extent, the bias would have to be implausibly large to completely undermine the qualitative conclusion of systematic underpricing.

4.5. *Decomposing supply-side residuals.* It is useful to relate supply-side residuals from the model to marginal cost estimates from the data with the following equation:

$$r_{jmt} = \hat{c}_{jmt} + \hat{d}_{jmt} \tag{17}$$

where $\hat{d}_{jmt} \equiv r_{jmt} - \hat{c}_{jmt}$ corresponds to the *distance* between marginal revenue and marginal cost, given the demand system, and observed prices, market shares, and costs. Alternatively, $\hat{d}_{jmt}$ can be interpreted as capturing a firm's *deviation* from the conduct assumption of profit maximization. In what follows, I will call this the "conduct deviation" term. When using differentiated products models to identify marginal costs, researchers implicitly assume conduct deviations of $\hat{d}_{jmt} = 0$; they presume their conduct assumption is correct and that supply-side residuals exactly correspond firms' marginal costs. The model validation results strongly suggest this assumption is invalid in the current setting.

With the content and non-content marginal costs components of $\hat{c}_{jmt}$, I can estimate the marginal cost functions in equations (6a) and (6b). With the $\hat{d}_{jmt}$ values, I can estimate models that allow me to investigate whether certain firm or market characteristics predict conduct deviations.[49] As Section 5 will show, both of these models are critical for predicting and evaluating the impact of cable mergers.

4.5.1. *Marginal cost functions.* Using the empirical analogues to content and non-content marginal costs from equations (14a)-(14d), I estimate the marginal cost functions in (6a) and (6b) for basic and non-basic cable by OLS. I use the average variable cost-based estimates of content and non-content marginal costs and not their translog counterparts for three reasons. First, average variable costs exhibit a much stronger correlation with the supply-side residuals than the translog-based marginal costs. Firms appear to

[49]The conduct deviation term can similarly be broken down into its "content" and "non-content" components. These effectively capture deviations from the channel count and pricing first order conditions. As the validation results from Table 6 foreshadow, I find conduct deviations and empirical relationships are mainly driven by the non-content component (or the pricing first order condition). In the interest of brevity, I only discuss results from regressions based on the total conduct deviation values.

set prices based on the former and not the latter. Second, flexibly modeling the content and non-content components of marginal costs as a function of firm and market characteristics with a translog cost function would require me to estimate third-order polynomials. Doing so is prohibitive given my sample size. In contrast, estimating marginal cost functions based on the average-variable cost components is straightforward. Third, using standard calculations based on the estimated translog cost functions, I find, like Kelly and Ying (2003), that the estimated long-run marginal cost functions suggest constant returns empirically.[50] This result lends support my use of average variable costs as proxies for marginal costs, as well as the common assumption in the literature that costs scale with the number of subscribers in a cable system.

Table 7 contains the estimation results. Panel A contains the marginal cost function parameter estimates that are based on the cost data. Panel B reports analogous estimates based on inferred marginal costs from the structural model. A quick comparison across the two panels reveals severe biases in the latter set of parameter estimates. This further illustrates the unreliability of the supply-side model for identifying the underlying cost structure of the industry.

[TABLE 7 ABOUT HERE]

The more reliable Panel A estimates show that larger cable companies have lower content and non-content costs for both basic and non-basic cable. These findings are consistent with larger firms: (1) being able to negotiate lower per-subscriber channel affiliation payments; and (2) realizing cost efficiencies through better management of local non-content labor and operating costs in their cable systems. MSOs similarly realize non-content cost efficiencies as compared to single-system operators, yet have relatively higher content costs. This is potentially due to MSOs offering relatively higher per-channel quality for basic and non-basic cable compared to their single-system counterparts. The firm fixed effects are mixed, indicating that beyond firm size and MSO effects, there is important unobserved heterogeneity in firms' costs. The remaining cost coefficients have their expected signs. For instance, per-subscriber content and non-content costs tend to

[50]The specific calculation from the translog function is the total cost elasticity of scale. This is computed using the estimated quantity coefficients from equation (15): $E_Q = \frac{1}{\hat{\rho}_1 + \hat{\rho}_2}$. For markets with both basic and non-basic cable I find $E_Q = \frac{1}{0.552+0.440} = 1.008$ at the sample mean. This implies that for a typical cable system a 1% increase in cable system size reduces per-unit costs by only 0.008%. Kelly and Ying (2003) estimate a similar translog cost function with similar cable system cost data and also find a small output elasticity of $E_Q = 1.03$. In markets with only basic cable, the output elasticity is higher at $E_Q = \frac{1}{0.918} = 1.09$, which suggests larger, yet still modest scale economies. These parameter estimates are listed in Table B.1 of online Appendix B.

be lower in large urban cable systems where agglomeration economies reduce the per-household cost of providing cable services.

To quantify the magnitude of the overall cost (in)efficiencies implied by the firm size, MSO, and firm fixed effects, I consider a set of hypothetical acquisitions in which an MSO acquires the licenses of all the single system cable operators in the sample. For a given set of acquisitions, I compute the predicted content and non-content costs under two scenarios: one where the single system operators own these licenses (i.e., what is observed), and another where one of the large MSOs owns these licenses. I then compute the percentage change in the predicted costs for each license under these two scenarios. This yields a distribution of cost (in)efficiencies across the acquired licenses for a given MSO.

Figure 3 plots the median and inter-quartile range of these distributions for content and non-content costs assuming different acquiring MSOs. A few key patterns emerge. Panel A highlights large non-basic content costs savings that are typically in the range of 25%, while basic content cost savings tend to be small. This may not be surprising given that basic cable is mainly comprised of broadcast networks that do not entail affiliate fees (unlike non-basic channels). This implies that dominant firms' negotiating advantage is relatively muted when it comes to basic content costs. In contrast, Panel B shows that content cost efficiencies tend to work through basic cable services. The non-content cost savings are quite large, in the range of 40%-50% for the dominant four MSOs (Rogers, Shaw, Cogeco, Vidéotron), which suggests that the industrial consolidation process would have resulted in large labor and operating cost efficiencies. Moreover, given basic non-content costs are incurred for both basic and non-basic cable because of the tying requirement, these cost efficiencies likely have an indirect effect on basic and non-basic prices over time as the industry consolidated. I return to this issue below.

[FIGURE 3 ABOUT HERE]

4.5.2. *Conduct functions.* To examine how conduct deviations vary across products, firms, and markets I run the following regression:

$$\hat{d}_{jmt} = 1\{Class^1_{mt}\}\delta_{1j} + 1\{Class^2_{mt}\}\delta_{2j} + w_{1mt}\eta_j + \zeta_{jmt} \tag{18}$$

I call these equations "conduct functions" since they reflect departures from the assumption of profit maximization. As with equation (12), the δ_{1j} and δ_{2j} estimates will correspond to systematic differences between marginal revenue and costs due to basic price regulation and the requirement that basic services must be included in the non-basic package. I purposely use w_{jmt} as the set of explanatory variables so I can assess conduct-related biases in the marginal cost function estimates based on the supply-side model's implied marginal costs.

The estimation results, which are presented in Table 8, yield a number of interesting findings. The Class 1 and Class 2 coefficients highlight how price regulation drives a wedge between marginal revenue and cost for larger, price-regulated markets. Moreover, the relative magnitudes of the coefficients are consistent with the CRTC's policy of enforcing more stringent price regulation in Class 1 licenses. The similarity of the coefficients in the "Basic" and "Non-Basic" columns indicates that the non-basic conduct deviations are mainly driven by basic cable conduct deviations, which necessarily impact the non-basic first order condition because of the tying requirement.

[TABLE 8 ABOUT HERE]

The estimates further indicate that the nationally dominant MSOs, Rogers, Shaw, and Cogeco, price much closer to the first order conditions than smaller cable companies. For example, their basic cable coefficient estimates of \$10.98, \$10.11 and \$17.76 compare to a sample median basic conduct deviation of -\$16.09. This suggests that these MSOs close the gap between marginal revenue and costs when setting their prices. The non-basic regression coefficients for these companies are also large compared to the median non-basic conduct deviation of \$18.36. Overall, these results suggest that the standard differentiated products model better approximates the behavior of large firms with a national presence than small locally-owned companies.[51]

In contrast, I find Vidéotron does price as aggressively as Rogers, Shaw, and Cogeco. This highlights the presence of important company-specific differences in conduct. These may be driven by cultural differences in the French and English parts of the country as Vidéotron is a Québec-based MSO.

There are two additional findings of note. First, cable operators in larger, richer markets tend to price closer to the first order conditions. Second, basic cable conduct deviations tend to be smaller in markets with only basic cable services. That is, all else being equal, companies appear to make smaller optimization errors for basic cable when only one product is offered.

[51]Large, firm-specific differences in conduct deviations also exist for other large MSOs in the sample such as Trillium or MacLean-Hunter.

There are a number of potential explanations for these patterns. The fact that major MSOs make substantially smaller optimization errors could reflect differences in expertise and knowledge across firms. For instance, nationally dominant firms may be able to afford to hire revenue managers to determine their profit-maximizing prices, while smaller firms cannot.

The finding that firms in less populous markets tend to set excessively low prices that result in large conduct deviations could reflect alternative objectives other than profit maximization. Indeed, in the sample's rural markets there are various community-run cable cooperatives as well as government-run cable systems (such as those on military bases) where the local cable operator primarily set prices to cover costs, and not to strictly maximize profits.

That optimization errors are larger in smaller cable systems could also reflect the fact that firms in these markets have smaller samples of subscribers with which to learn about local cable demand. This is in contrast to city-based cable systems, where firms have large samples of households to estimate demand elasticity. If firms in smaller markets are indeed more uncertain about local demand, and further if they tend to over-estimate demand elasticity, then this could explain the finding that conduct deviations are more negative in smaller markets.[52]

I stop short of incorporating additional structure into the model to quantify the impact of these and other mechanisms in generating conduct deviations; doing so is beyond the scope of the current article. Regardless of underlying mechanism(s), the existence of systematic conduct deviations has broader implications for the use of structural market demand/supply models in measuring firms' marginal costs and informing policy. In the next section, I consider an application that has had a particularly large policy impact among anti-trust authorities worldwide: merger simulations.

5. EVALUATING THE IMPACT OF CABLE CONSOLIDATION

This section evaluates the impact of cable mergers on prices, cable bundles, costs, and welfare. Keeping with a recent trend in the literature (Peters 2006; Weinberg and Hosken 2012; Houde 2012; Bjornerstedt and Verboven 2012), I provide comparative structural and reduced-form analyses of merger effects. By comparing simulated merger effects based on structural model to estimated effects based on actual mergers

[52]A similar sample size-based intuition could also explain why major MSOs make substantially smaller optimization errors than smaller cable companies. MSOs serve large samples of subscribers across the country which they can use to obtain accurate estimates of cable demand. Such demand estimates can in turn be used to identify profit-maximizing price levels.

in the data, I can evaluate the model's performance in forecasting merger outcomes.

Methodologically, I build on the standard merger simulation approach by incorporating cost data into the analysis. Unlike existing studies, I account for conduct specification errors in evaluating merger outcomes. Through the analysis I also provide some of the first estimates of the impact that cable TV consolidation has on cable prices and bundles, consumer welfare, and firms' profits.

5.1. *Retrospective merger simulations.* Using the structural model, I quantify merger effects by simulating market outcomes under two scenarios: one with consolidation (observed in the data) and another without (the counterfactual). The latter "no-consolidation" scenario assumes that each license's cable company is what is was in 1990 for the entire sample period.

Two sets of simulations are used. The first corresponds directly to the standard merger simulation approach. For these simulations, I assume firms' content and non-content marginal costs are the inferred values of $c^x_{1mt}, c^x_{2mt}, c^{nx}_{1mt}, c^{nx}_{2mt}$ from the first order conditions in (8a) and (8b). With these inferred costs, I estimate the content and non-content marginal cost functions by running the regressions (10a) and (10b) for basic and non-basic cable. The resulting marginal cost function parameter estimates are reported in Panel B of Table 7.

Using the estimated regression coefficients and residuals, I predict content and non-content costs for package j under scenario $s \in \{0, 1\}$:

$$c^{x,s}_{jmt} = w^s_{jmt}\tilde{\gamma}^x_j + \tilde{\omega}^x_{jmt} \tag{19a}$$

$$c^{nx,s}_{jmt} = w^s_{jmt}\tilde{\gamma}^{nx}_j + \tilde{\omega}^{nx}_{jmt} \tag{19b}$$

where $\tilde{\gamma}^x_j$ and $\tilde{\gamma}^x_j$ are the regression coefficient estimates from Panel B of Table 7, and $\tilde{\omega}^x_{jmt}$ and $\tilde{\omega}^{nx}_{jmt}$ are the corresponding residuals from the marginal cost regressions. The s superscript in w^s_{jmt} indexes the cost shifters by scenario s. Under the $s = 0$ (consolidation) scenario, the predicted values correspond exactly to the observed content and non-content costs in the data: $c^{x,0}_{jmt} = c^x_{jmt}$ and $c^{nx,0}_{jmt} = c^{nx}_{jmt}$. This follows standard practice in the literature on merger simulations: the residuals from the marginal cost regressions are assumed to be policy-invariant structural cost shocks that firms account for in making their pricing and bundling decisions. The cost predictions under the $s = 1$ (no consolidation) scenario will differ because of the consolidation-induced changes in the $FirmSize_{mt}$, MSO_{mt} and $Firm_{mt}$ variables in the w^s_{jmt} vector.

Cable programming quality will also differ under the two scenarios. Recall that MSO_{mt} and $Firm_{mt}$ are included in y^0_{jmt} in equation (1) (which affects non-content cable quality/branding effects), and that the MSO_{mt} and $LargeFirm_{mt}$ variables are in z_{mt} in equations (2a) and (2b) (which affect per-channel programming content quality). Therefore, to predict cable prices and channel counts with and without consolidation, I solve the first order conditions in (8a) and (8b) using the cost predictions for $\tilde{c}^{x,s}_{jmt}$ and $\tilde{c}^{nx,s}_{jmt}$, together with their corresponding utility function variables in $y^{0,s}_{jmt}$ z^s_{mt} for scenarios $s \in \{0,1\}$.

In light of the results from Section 4.4, the structural model will yield unreliable estimates of any merger-induced cost (in)efficiencies. This is because the inferred costs from the supply-side model are affected by firms' actual marginal costs *and* any deviations from the first order conditions in (8a) and (8b). This does not imply, however, that standard merger simulations will necessarily yield inaccurate forecasts of merger effects on prices and channel counts. The above results suggest a more nuanced interpretation of the equations in (8a) and (8b) that are used to generate these forecasts: they are a nonlinear system of equations that are based on a differentiated products model, but that incorporates systematic deviations from the model's first order conditions. In simulating prices and channel counts with and without consolidation, the difference in the predicted outcomes not only reflect merger-induced changes in firms' costs, but also their conduct.

To better understand the issues involved in interpreting the standard merger simulation results in the presence of conduct deviations, consider a hypothetical acquisition of a small, locally-owned firm's licenses by the largest MSO in the sample, Rogers. The predicted cost efficiencies from Figure 3 suggest that Rogers will reduce content and non-content costs in the acquired licenses, allowing it to potentially offer more channels and charge lower prices. The conduct function estimates from Table 8 indicate that Rogers will, however, tend to set prices more in-line with the first order conditions in (8a) than the small firm previously did. If the small firm was previously setting prices well below profit maximizing levels (which recall from Figure 2 is typically the case), then the differences in conduct between the two firms will see Rogers increase prices until marginal revenues are closer to the (new) marginal costs. The simulated merger effects based on the first order conditions in (8a) (8b), and the predicted marginal costs under scenarios $s = 0,1$ from (19a) and (19b), will capture the net effect of the cost- and conduct-related forces in driving merger outcomes. This directly follows from the relationship between the model's inferred marginal costs, the estimated marginal costs from the cost data, and conduct deviations in (17).

The second set of simulations serve as a benchmark for the standard merger simulation results. For these simulations, I use the empirical analogues of the marginal costs from equations (14a)-(14d), and marginal cost function estimates from Panel A of Table 7. With the latter set of estimates I can similarly predict marginal costs with and without consolidation. I again treat the residuals from the content and non-content cost regressions as structural cost shocks in doing so. Using these two sets of marginal costs, I simulate merger outcomes under the two scenarios by solving the constrained profit maximization problem in (7) under the consolidation and no consolidation scenarios. Importantly, these simulations also make use of the basic cable price caps from the CRTC Master Files in solving for firms' constrained profit-maximizing cable prices and channel counts.

The advantage of the second set of simulations is that they will yield more reliable estimates of the impact mergers have on costs. This is because they use direct cost measures that do not rely on any assumptions regarding firms' conduct. The shortcoming of these simulations is they rely on a specific conduct assumption (constrained profit maximization) to generate predictions over prices and channel counts with and without consolidation. Unlike the standard merger simulations based on the nonlinear system of equations in (8a) and (8b), these simulations will not incorporate firm-specific differences in conduct when predicting merger effects. To the extent that differences in conduct matter quantitatively across acquiring and acquired firms, these simulations will exhibit forecast errors when predicting the impact cable mergers have on prices and channel counts.

For a given set of simulated outcomes, I construct a "difference-in-difference" estimate of the a merger effect in license m as follows. Suppose license m's cable operator changed in year τ_m because of a merger. The "difference" estimate of the merger's effect on basic cable prices (for example) under the consolidation scenario is the difference between the average basic price in years greater than or equal τ_m and the average basic price in years prior to τ_m. Denote this difference by Δp^0_{1m}. The difference-in-difference estimate of the merger on basic prices can then be computed as $\Delta\Delta p_{1m} \equiv \Delta p^0_{1m} - \Delta p^1_{1m}$, where Δp^1_{1m} is an analogous difference in average prices before and after the merger under the no-consolidation scenario. Similar calculations are used to estimate the causal impact of mergers on the other outcome variables of interest.

5.2. *Reduced-form estimates of merger effects.* To compare the predictive ability of the two sets of merger simulations just discussed, I require an empirical benchmark of merger effects. Following observational studies of historical mergers (for example, Dafny 2009 or Sweeting 2010), I use the following regression model to obtain a benchmark:

$$y_{jmt} = \alpha A_{mt} + \mathbf{X}_{jmt}\beta + \mu_m + \tau_t + v_{jmt} \tag{20}$$

where y_{jmt} is the outcome variable of interest: $y_{jmt} \in \{p_{jmt}, x_{jmt}, \hat{c}^{x}_{jmt}, \hat{c}^{nx}_{jmt}\}$ for $j = 1,2$.[53] The variable A_{mt} equals one if there has been a cable merger in license m during year t, or any year prior to t. The vector of controls, $\mathbf{X}_{jmt}$, includes average household income, urban density, system size, and the total kilometers of cable in a cable system. The model also includes license and year fixed effects, μ_m and τ_t, and an econometric error, v_{jmt}. The coefficient of interest, α, is a difference-in-difference estimate of the within-license impact of a merger. Using the language from experiments, licenses that experience a merger are in the "treatment group", while all other licenses are in the "control".

It is well-known that various econometric issues can undermine the identification of α. For example, α will be biased if the outcome variable has different trends across licenses in treatment and control. As a check on this issue, I examine whether year-to-year trends in cable prices, channel counts, and costs differ across the treatment and control groups prior to mergers. The corresponding (unreported) regression results that check this issue suggest no significant differences in pre-merger trends.

Perhaps more problematic is the reality that mergers are likely not randomly assigned to licenses, which would generate selection bias in the reduced-form merger effects estimates. Selection effects would have to work through the transitory component of a market's idiosyncratic unobservable, v_{jmt}, because the market and year fixed effects account for any license- or year-specific permanent unobserved demand or cost shocks. Given that a large fraction of acquisitions in the sample involve nationally dominant companies who tend to make merger decisions at the firm and not local level, it unclear whether selection based on these local idiosyncratic shocks will have a large impact on the estimates.

To the extent that selection effects exist, they will lead me to overstate merger effects for costs and bundle quality, while their impacts in price regressions are ambiguous. Suppose acquiring firms target licenses with large idiosyncratic cost shocks because there's opportunity to cut costs and realize efficiency gains. This selection effect would cause me to overstate any estimated cost efficiencies from acquisitions. Suppose

[53]Given the large variance of the dependent variables, as a robustness check I report results from analogous regressions with the dependent variable in logs in Table A.3 of online Appendix A. The signs and magnitudes of the estimated merger effects from these regressions largely correspond to the estimation results from equation (20).

instead that acquiring firms target licenses with high idiosyncratic cable quality demand shocks in order to improve local cable packages, offer more channels, and expand the market. This would also create an upward bias in α in the channel count regressions. In the price regressions, the selection bias would be negative under the cost shock story (since mark-ups and price levels depend on costs), while it would be positive under the demand-shock story (since firms can charge higher prices if they offer higher quality bundles). The direction of the bias in the price regressions is therefore ambiguous; it depends whether transitory demand or cost shocks dominate in determining firms' acquisition targets.

5.3. *Findings.* The reduced-form and simulation-based estimates of merger effects are presented in Table 9.[54] The difference-in-difference estimates indicate that following a cable merger, the acquiring firm lowers monthly cable prices by \$0.87, and add two more channels to the non-basic bundle. Basic channel counts likely remain unchanged because basic bundles mainly consist of local programming and national broadcast channels. Recall that both MSOs and small cable companies must offer these channels under the CRTC's must-carry regulations. This implies that in contrast to non-basic services, there is little scope for larger firms to expand upon the basic cable bundles of acquired smaller firms.

[TABLE 9 ABOUT HERE]

The estimates in the bottom panel of the table reflects the cost-reducing firm size and firm-specific effects from Panel A of Table 7. Upon entering new markets through acquisition, dominant firms reduce monthly per-subscriber basic and non-basic non-content costs by \$2.19 and \$1.63. These large 15% to 20% merger-related labor and operating cost efficiencies likely contribute to the merged-induced fall in basic cable prices. They may also explain the additional finding that non-basic prices remain unchanged following acquisitions despite the fact that more non-basic channels tend to be offered.

In contrast, the reduced-form estimates reveal relatively small or statistically insignificant changes in basic and non-basic content costs following cable mergers. This likely reflects two opposing forces. On the one hand, the content-cost efficiencies of larger MSOs should see these costs fall when larger acquiring firms enter new markets.[55] On the other hand, as the demand estimates and WTP results from Table 4 show, larger MSOs tend to offer higher per-channel cable quality which, all else equal, push up content

[54]The estimates are based on 142 instances where a market's cable company changed because of a merger in the sample.

[55]These cost efficiencies are implied by the marginal cost estimates in Panel A of Table 7, and are depicted in Figure 3.

costs.[56] The reduced-form merger effects estimates therefore identify the net effect of these merger-related cost efficiencies and improved per-channel cable quality.

The second and third columns of Table 9 present the predicted merger effects from the standard and modified merger simulations. The results in Panel A show that the standard merger simulations are indeed able to predict that basic cable prices fall while non-basic channel counts rise around mergers. In contrast, the modified simulations provide poor predictions of the impact of mergers on cable prices and channel counts. The bottom panel, however, shows that the modified simulations do a much better job of predicting merger-related cost changes than the standard simulations do. This is not surprising given the biases inherent to the structural model's marginal cost estimates, and the fact that the cost predictions for the modified simulations are based on the actual cost data.

5.4. *Welfare effects.* An additional benefit from having estimated a structural model is I can use it to quantify the impact of cable mergers on consumer welfare and profits. Following Nevo (2000) (and many others), I use compensating variation to measure the change in consumer welfare due to the consolidation process. This welfare measure is the amount of income a consumer would have to be paid under the no-consolidation scenario such that they would be indifferent between living in worlds with and without consolidation.

When estimating the consumer welfare effects of cable mergers, I use predicted prices and channel counts from the standard merger simulations because they yield better forecasts of the impact of mergers on these variables. The predicted costs from these simulations are, of course, severely biased and unreliable. As such, they are not useful for measuring the impact of mergers on profits. Therefore, to evaluate the impact that mergers have on consumer welfare *and* profits, I augment the standard merger simulations with the estimated marginal cost functions based on the cost data.

Specifically, I assume that the content and non-content marginal cost functions that are based on the cost

[56]More specifically, higher quality (i.e., more popular channels with better ratings) channels tend to demand larger per-subscriber affiliation payments from cable companies. See Crawford and Yurukoglu (2012) for direct evidence of this.

data and inferred costs (i.e., from Panels A and B of Table 7) are related as follows:

$$\begin{aligned} c^{x,s}_{jmt} &= w^{s}_{jmt}\tilde{\gamma}^{x}_{j} + \tilde{\omega}^{x}_{jmt} \\ &= \underbrace{w^{s}_{jmt}\hat{\gamma}^{x}_{j} + \hat{\omega}^{x}_{jmt}}_{\hat{c}^{x,s}_{jmt}} + \underbrace{w^{s}_{jmt}\hat{\delta}^{x}_{j} + \hat{\zeta}^{x}_{jmt}}_{\hat{d}^{x}_{jmt}} \end{aligned} \tag{21a}$$

$$\begin{aligned} c^{nx,s}_{jmt} &= w^{s}_{jmt}\tilde{\gamma}^{nx}_{j} + \tilde{\omega}^{nx}_{jmt} \\ &= \underbrace{w^{s}_{jmt}\hat{\gamma}^{nx}_{j} + \hat{\omega}^{nx}_{jmt}}_{\hat{c}^{nx,s}_{jmt}} + \underbrace{w^{s}_{jmt}\hat{\delta}^{nx}_{j} + \hat{\zeta}^{nx}_{jmt}}_{\hat{d}^{nx}_{jmt}} \end{aligned} \tag{21b}$$

where $\hat{d}^{x}_{jmt} \equiv c^{x}_{jmt} - \hat{c}^{x}_{jmt}$ and $\hat{d}^{nx}_{jmt} \equiv c^{nx}_{jmt} - \hat{c}^{nx}_{jmt}$ are the content and non-content components of conduct deviations for package j, and where $\tilde{\delta}^{x}_{j}, \tilde{\delta}^{nx}_{j}$ and $\tilde{\zeta}^{x}_{jmt}, \tilde{\zeta}^{nx}_{jmt}$ are the regression coefficient estimates and residuals from regressions of $\hat{d}^{x}_{jmt}$ and $\hat{d}^{nx}_{jmt}$ on w_{jmt}.[57] These equations make clear the relationships between the cost parameters and cost shocks based on the inferred costs and cost data: $\tilde{\gamma}^{x}_{j} = \hat{\gamma}^{x}_{j} + \hat{\delta}^{x}_{j}$ and $\tilde{\omega}^{x}_{jmt} = \hat{\omega}^{x}_{jmt} + \hat{\zeta}^{x}_{jmt}$ (and similarly for the non-content cost equations).

In short, I use the estimated cost functions based on the cost data to separately identify content costs and non-content costs and conduct deviations in the model's inferred costs. By augmenting the standard merger simulation framework with cost data, I am able to predict the impact of cable mergers on prices, costs, profits, and consumer welfare.

By defining the model's inferred costs according to (21a) and (21b), I am also able to quantify the impact of cost-reducing firm size effects in generating the welfare gains/losses from consolidation. Consider, for example, the impact of firm-size related cost efficiencies in non-content costs generated by mergers. Recall that I simulate market outcomes and welfare under the consolidation scenario by setting the cost, non-content quality, and per-channel programming content quality variables in w^{s}_{jmt}, $y^{0,s}_{jmt}$, and z^{s}_{jmt} to their $s = 0$ values. To "turn off" the cost-reducing firm size effects on non-content costs, I start from these consolidation variables, and change the values for $FirmSize_{mt}$ in the w^{s}_{jmt} vector to their $s = 1$ values in the $\hat{c}^{nx,s}_{jmt}$ (and hence $c^{nx,s}_{jmt}$) functions in (21b). That is, I use this w^{s}_{jmt} vector, to generate "partial" no consolidation values for $\hat{c}^{nx,s}_{jmt}$ and $c^{nx,s}_{jmt}$. I then simulate cable prices and channel counts using these (new) $c^{nx,s}_{jmt}$ values in conjunction

[57]It write out the conduct deviation functions in (21a) and (21b) to help illustrate how both costs and conduct deviations affect the implied content and non-content marginal costs from the model. In the simulations below, I do not explicitly use the conduct deviation functions to investigate merger effects, and therefore do not report $\tilde{\delta}^{x}_{j}$ $\tilde{\delta}^{nx}_{j}$.

with the consolidation scenario $c^{x,s}_{jmt}$ values and the first order conditions in (8a) and (8b). By comparing consumer and producer welfare under this "partial" no-consolidation scenario to the consolidation scenario, I am able to quantify the role of non-content cost-reducing firm size effects in driving the predicted welfare gains/losses from consolidation.[58]

There are some important caveats to bear in mind when interpreting the welfare results. Taken literally, the model predicts that consumers immediately receive a utility "bump" in terms of programming content quality when dominant MSOs acquire their local cable providers. This may overstate the welfare gains from consolidation if it takes longer than a year for dominant cable companies to develop the infrastructure of acquired cable systems to improve local programming.

Moreover, because I do not model firms' bundling decisions year-to-year, I am unable to say whether the introduction of certain channels or the re-bundling of existing channels (i.e., leaving bundle size unchanged) by dominant acquiring firms are the key drivers of any programming-related welfare gains. The model approximates the welfare gains from this deeper bundling process by giving the same MSO-specific utility "bump" to both existing channels and new channels in a license's basic and non-basic bundles following an acquisition.

I also abstract from the impact of cable consolidation on upstream channel providers. The negotiated affiliation payments that channel companies receive likely fall as dominant cable providers take over the industry, leaving them worse off. In the long-run, this could cause content providers to exit, leaving consumers with less programming variety. The short-run merger analysis abstracts from these interesting long-run effects of cable consolidation. Incorporating long-run effects into the analysis is beyond the scope of this paper and is left for future work.

5.4.1. *Results.* The predicted welfare effects are presented in Table 10. I again report "difference-in-difference" estimates of welfare effects based on the simulated merger outcomes. More specifically, for each market that experiences a merger, I compute the expected consumer surplus and profits per subscriber in the years before and after the merger. These are computed under the consolidation and no consolidation scenarios. I then compute the difference in expected consumer surplus/profits before after the merger under

[58] I focus on non-content cost reducing firm size effects because: (1) the difference-in-difference estimates highlight statistically significant and large reductions in non-content costs following a cable merger; and (2) cost-reducing firm size effects in the estimated non-content marginal costs for basic cable are statistically significant and large in Panel A of Table 7.

these two scenarios. Finally, I compute the difference in these differences to quantify the welfare impact of consolidation. The table reports the sample averages and standard deviations of the differences in differences across all markets that were affected by a cable merger.

[TABLE 10 ABOUT HERE]

The main takeaway from the table is that both consumer welfare and firms' profits rise as a result of industrial consolidation. On average, consumers realize an additional $2.19 of surplus of month following acquisitions, while firms realize an additional $3.63 of profits. To provide a sense of the magnitude of these welfare changes, they are roughly 10% and 15% of the sample average for basic cable prices.

The bottom part of the table illustrates the importance of cost-reducing firm size effects in non-content costs in driving these consolidation-induced welfare effects. In the absence of cost-reducing firm size effects, the average gains in consumer surplus and profits from consolidation become statistically insignificant, falling to $0.76 and $0.84. Intuitively, dominant acquiring firms' non-content cost efficiencies play an important role in keeping cable prices down following mergers, which directly benefits consumers.

6. SUMMARY AND DISCUSSION

This paper has provided a novel evaluation of the benchmark differentiated products model from empirical IO, focusing on its ability to identify the actual marginal costs of firms. Using unique demand and cost data from the cable TV industry in Canada, I found evidence that the model performs poorly in doing so, and that the biases in the model's marginal cost estimates are largely due to firms pricing below profit maximizing levels. I studied the implications for merger simulations, and found these biases compromise estimates of merger-related cost efficiencies, but did not prevent the model from generating useful predictions of the price and non-price effects of mergers.

Through the analysis, I also provided novel estimates of the impact of consolidation on cable prices, channels, costs, and welfare. I found that dominant cable companies reduce basic cable prices, increase the number of non-basic channels, and reduce per-subscriber costs when entering new cable markets through acquisition. The net effect of these changes is that both consumer welfare and firms' profits tend to rise following a merger.

These results have implications for policymakers and research on differentiated products markets. In terms of policy, the findings are relevant for anti-trust authorities who actively use merger simulations to evaluate prospective mergers. A key policy implication is that authorities should use their power of subpoena to request cost data from firms to enhance the accuracy and transparency of their merger simulations.[59] The paper not only motivates such requests, it is prescriptive as it shows how an analyst can augment the standard merger simulation methodology with cost data.

The model validation results have implications for empirical research on differentiated products markets more broadly. They serve as a warning about the effects of conduct misspecification on the implied marginal costs from differentiated products models. Given the severity of these errors, the paper's results suggest that if cost data is unavailable and conduct cannot be identified in a differentiated product model (as is typically the case), researchers should, at the very least, produce results based on different assumptions over the form of conduct. These might include Bertrand, collusion, or partial coordination among some firms. Further, if "natural experiments" such as mergers or other policy shocks exist, they can be exploited to assess the predictive ability of the model under different conduct assumptions to validate the supply-side of the model.

The results have at least three implications for empirical research on the cable TV industry. Most papers in the literature have adopted the conduct assumption of profit maximization to identify firms' marginal costs and evaluate policy (for example, Rubinovitz 1993 or Chu 2010). My results serve as a caveat to these papers' results. Second, the results lend direct support to the estimation strategy of Crawford and Yurukoglu (2012), who go beyond inverting first order conditions to recover firms' marginal costs, and identify costs by using moments based on national-average input cost data for individual channels. By exploiting cost-based moments in this way, future studies can mitigate the impact of conduct misspecification in identifying the industry's cost structure. Finally, my finding that differentiated products models better approximate the behavior of dominant cable companies and companies that operate in larger markets suggest that future studies should focus on these types of firms and markets when estimating equilibrium models for the industry. Luckily, because of the consolidation process studied in this paper, many cable markets today are served by such dominant firms. This implies that the profit maximization assumption should be more reliable with more contemporary datasets.[60]

[59]For example, the U.S. Hart-Scott-Rodino Act (1976) gives the Department of Justice and the Federal Trade Commission such power to issue a "second request" for data from parties who are proposing to merge.

[60]The same cannot be said, however, for early studies in the literature by Otsuka and Mayo (1991) and Rubinovitz (1993) on the impact of basic price regulation on cable prices. These authors also assume strict profit maximization, and use demand-side data from the 1990s that is similar to what I use. In light of my findings, it is possible that these papers overstate the effectiveness

To close, I highlight a few avenues for future research. While studying cost and conduct misspecification in the context of local monopolies has been convenient, it would be useful to undertake a similar study in an oligopoly context. A key question of interest is whether firms hold beliefs/conjectures about competition that correspond to the Bertrand assumption commonly used in the literature. More generally, future research that studies on how cost data can be used to develop and identify differentiated products models whose predictions are robust to conduct misspecification should prove valuable.[61]

Regarding the impact of cable TV consolidation, the short-run analysis in this paper has ignored the impact of downstream consolidation among cable companies upstream channel providers. Going forward, it would be useful to build on the structural model, and examine the extent to which channel providers have an incentive to exit as a result of receiving smaller negotiated affiliation payments from increasingly large cable companies.

Department of Economics, The University of Melbourne, Australia

REFERENCES

BERRY, S. T., J. LEVINSOHN AND A. PAKES, "Automobile Prices in Market Equilibrium," Econometrica 63 (1995), 841–890.

BERRY, S. T. AND J. WALDFOGEL, "Do Mergers Increase Product Variety? Evidence from Radio Broadcasting," Quarterly Journal of Economics 55 (2001), 49–67.

BJORNERSTEDT, J. AND F. VERBOVEN, "Does Merger Simulation Work? A "Natural Experiment" in the Swedish Analgesics Market," mimeo, Catholic University of Leuven, 2012.

BRESNAHAN, T., "The Oligopoly Solution Concept is Identified," Economics Letters 10 (1982), 87–92.

———, "Competition and Collusion in the American Automobile Industry: The 1955 Price War," Journal of Industrial Economics 35 (1987), 457–482.

of price regulation in preventing monopoly pricing.

[61] In Byrne et al. (2013), we establish that the demand and cost parameters of differentiated products models can be jointly identified with demand-side data on prices, market shares, and product characteristics, some information on firms' total costs, the profit maximization assumption, and standard regularity conditions on cost functions. That is, we formalize how cost data can be used to help identify cost functions *and* demand systems. They key novelty of our approach is that we can identify all the model's parameters without resorting to instrumental variables. Our work is not unrelated to De Loecker (2011) who shows how production function sand productivity estimates can be improved with some information on the demand system.

———, “Empirical Studies of Industries with Market Power,” in R. Schmalensee and R. D. Willig, eds., Handbook of Industrial Organization Volume 2 (Amsterdam: North-Holland, 1989), 1011–1057.

BYRNE, D. P., “An Empirical Study of the Causes and Consequences of Mergers in the Canadian Cable Television Industry,” Ph.D. Dissertation, Queen’s University, 2010.

BYRNE, D. P., M. HIRUKAWA, S. IMAI AND V. SARAFIDIS, “Instrument-Free Identification and Estimation of Differentiated Products Models with Cost Data,” mimeo, The University of Melbourne, 2013.

CALFEE-STAHL, J., “A Dynamic Analysis of Consolidation in the Broadcast Television Industry,” mimeo, Board of Governors of the Federal Reserve System, 2011.

CHIPTY, T., “Horizontal Integration for Bargaining Power: Evidence From the Cable Television Industry,” Journal of Economics and Management Strategy 4 (1995), 375–397.

———, “Vertical Integration, Market Foreclosure, and Consumer Welfare in the Cable Television Industry,” American Economic Review 91 (2001), 428–453.

CHIPTY, T. AND C. M. SNYDER, “The Role of Firm Size in Bilateral Bargaining: A Study of the Cable Television Industry,” Review of Economics and Statistics 81 (1999), 326–340.

CHRISTENSEN, L., D. W. JORGENSON AND L. J. LAU, “Transcendental Logarithmic Production Frontiers,” Review of Economics and Statistics 55 (1973), 28–45.

CHU, C. S., “The Effect of Satellite Entry on Cable Television Prices and Product Quality,” RAND Journal of Economics 41 (2010), 730–764.

CHU, C. S., P. LESLIE AND A. SORENSEN, “Bundle-Size Pricing as an Approximation to Mixed Bundling,” American Economic Review 101 (2011), 263–303.

CLAY, K. AND W. TROESKEN, “Further Tests of Static Oligopoly Models: Whiskey, 1882-1898,” Journal of Industrial Economics 51 (2003), 151–166.

CRAWFORD, G., “The Impact of the 1992 Cable Act on Household Demand and Welfare,” RAND Journal of Economics 32 (2000), 422–449.

———, “Cable Regulation in the Internet Era,” in N. Rose, ed., Economic Regulation and Its Reform: What Have We Learned? (Chicago: Chicago University Press, 2013), Chapter 4.

CRAWFORD, G. AND M. SHUM, “The Welfare Effects of Endogenous Quality Choice: The Case of Cable Television,” mimeo, California Institute of Technology, 2006.

CRAWFORD, G. S. AND M. SHUM, "Monopoly Quality Degradation in Cable Television," Journal of Law and Economics 50 (2007), 181–219.

CRAWFORD, G. S. AND A. YURUKOGLU, "The Welfare Effects of Bundling in Multichannel Television Markets," American Economic Review 102 (2012), 301–317.

DAFNY, L., "Estimation and Identification of Merger Effects: An Application to Hospital Mergers," Journal of Law and Economics 52 (2009), 523–550.

DE LOECKER, J., "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," Econometrica 79 (2011), 1407–1451.

DRAGANSKA, M., M. MAZZEO AND K. SEIM, "Beyond Plain Vanilla: Modeling Joint Product Assortment and Pricing Decisions," Quantitative Marketing and Economics 7 (2009), 105–146.

DUBE, J.-P., J. T. FOX AND C.-L. SU, "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," Econometrica 80 (2012), 2231–2267.

FAN, Y., "Ownership Consolidation and Product Characteristics: A Study of the U.S. Daily Newspaper Market," American Economic Review 103 (2013), 1598–1628.

FCC, "Report on Cable Industry Prices," Technical Report, Federal Communications Commission, 2002.

FISHER, F. M. AND J. J. MCGOWAN, "On the Misuse of Accounting Rates of Return to Infer Monopoly Power," American Economic Review 73 (1983), 82–97.

FORD, G. S. AND J. D. JACKSON, "Horizontal Concentration and Vertical Integration in the Cable Television Industry," Review of Industrial Organization 12 (1997), 501–518.

GENESOVE, D. AND W. P. MULLIN, "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914," RAND Journal of Economics 29 (1998), 355–377.

GOLDBERG, P. K., "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," Econometrica 63 (1995), 891–951.

GOOLSBEE, A. AND A. PETRIN, "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV," Econometrica 72 (2004), 351–381.

HASTINGS, J. S., "Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California," American Economic Review 94 (2004), 317–328.

HORTASCU, A. AND S. L. PULLER, "Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market," RAND Journal of Economics 39 (2008), 86–114.

HOUDE, J. F., "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," American Economic Review 102 (2012), 2147–2182.

JEZIORSKI, P., "Effects of Mergers in Two-sided Markets: The U.S. Radio Industry," American Economic Journal: Microeconomics (2013a), forthcoming.

———, "Estimation of Cost Synergies from Mergers: Application to the U.S. Radio," RAND Journal of Economics (2013b), forthcoming.

KELLY, M. T. AND J. S. YING, "On Measuring Competitive Viability and Monopoly Power in Cable: An Empirical Cost Approach," Review of Economics and Statistics 85 (2003), 963–970.

KIM, D.-W. AND C. R. KNITTEL, "Biases in Static Oligopoly Models? Evidence from the California Electricity Market," Journal of Industrial Economics 54 (2003), 451–470.

KNITTEL, C. AND K. METAXOGLOU, "Estimation of Random Coefficient Demand Models: Challenges, Difficulties, and Warnings," Review of Economics and Statistics (2012), forthcoming.

MELITZ, M. J. AND S. J. REDDING, "Heterogeneous Firms and Trade," in G. Gopinath, E. Helpman and K. Rogoff, eds., Handbook of International Economic, Volume 4 (Oxford: North-Holland, 2014), 1–55.

NEVO, A., "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," RAND Journal of Economics 31 (2000), 395–421.

———, "Measuring Market Power in the Ready-to-Eat Cereal Industry," Econometrica 69 (2001), 307–342.

OTSUKA, Y. AND J. W. MAYO, "Demand, Pricing, and Regulation: Evidence from the Cable TV Industry," RAND Journal of Economics 22 (1991), 396–410.

PARSONS, P. R., "Horizontal Integration in the Cable Television Industry: History and Context," Journal of Media Economics 16 (2003), 23–40.

PETERS, C., "Evaluating the Performance of Merger Simulations: Evidence from the U.S. Airline Industry," Journal of Law and Economics 49 (2006), 627–649.

PETRIN, A., "Quantifying the Benefits of New Products: The Case of the Minivan," Journal of Political Economy 110 (2002), 705–729.

RUBINOVITZ, R. N., "Market Power and Price Increases for Basic Cable Service since Deregulation," RAND Journal of Economics 24 (1993), 1–18.

SLADE, M., "Merger Simulations of Unilateral Effects: What can we Learn from the UK Brewing Industry?," in B. Lyons, ed., Cases in Europeak Competition Policy, The Economic Analysis (Cambridge: Cambridge University Press, 2009), Chapter 13.

SMILEY, A., "Direct Competition Among Cable Television Systems," Economic Analysis Group Discussion Paper EAG 86-9, Antitrust Division - U.S. Department of Justice., 1986.

SMITH, H. AND C. O'GORMAN, "Efficiency Gain from Ownership Deregulation: Estimates for the Radio Industry," mimeo, Oxford University, 2008.

SWEETING, A., "The Effects of Mergers on Product Positioning: Evidence from the Music Radio Industry," RAND Journal of Economics 41 (2010), 372–397.

TYLER-MOONEY, C., "A Two-Sided Market Analysis of Radio Ownership Caps," mimeo, University of Oklahoma, 2010.

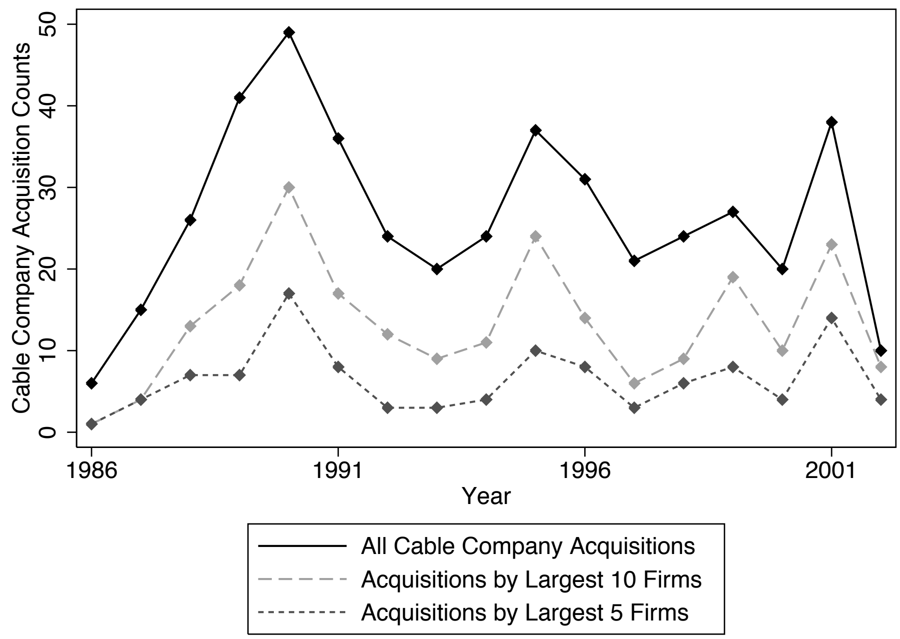
WEINBERG, M. AND D. HOSKEN, "Evidence on the Accuracy of Merger Simulations," Review of Economics and Statistics (2012), forthcoming.

WIETEN, J., G. MURDOCK AND P. DAHLGREN, Television Across Europe, A Comparative Introduction (London: SAGE Publications Ltd., 2000).

WOLFRAM, C., "Measuring Duopoly Power in the British Electricity Spot Market," American Economic Review 89 (1999), 805–826.

Figure 1: Trends in Cable Mergers and National Market Shares of Large Cable Companies

Panel A: Annual Merger Counts



Panel B: National Market Share of Dominant Cable Companies

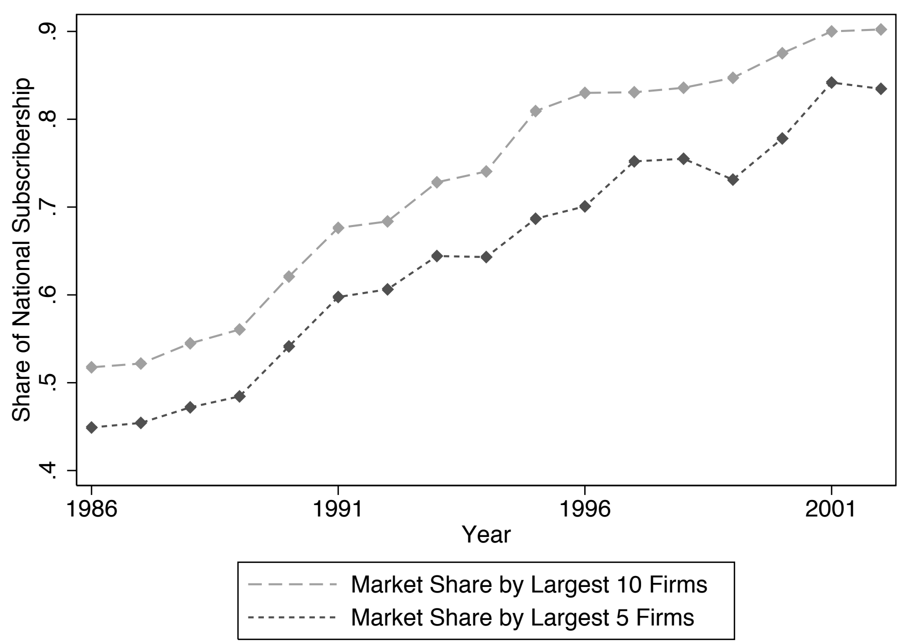


Figure 2: Predicted Marginal Revenue, Proft-maximization Error, and Demand Elasticity in Unregulated Cable TV Markets as a Function of the Utility Function Price Coefficient

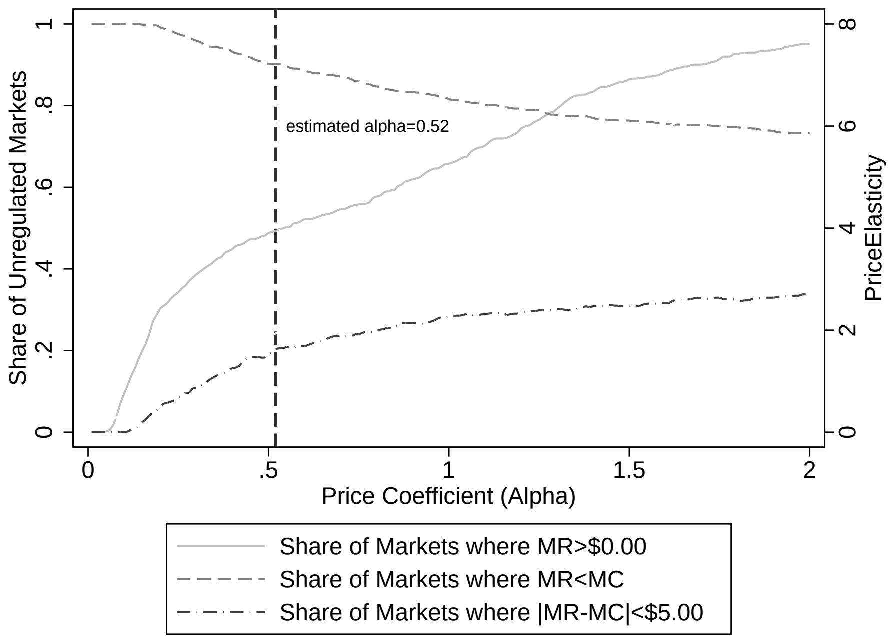
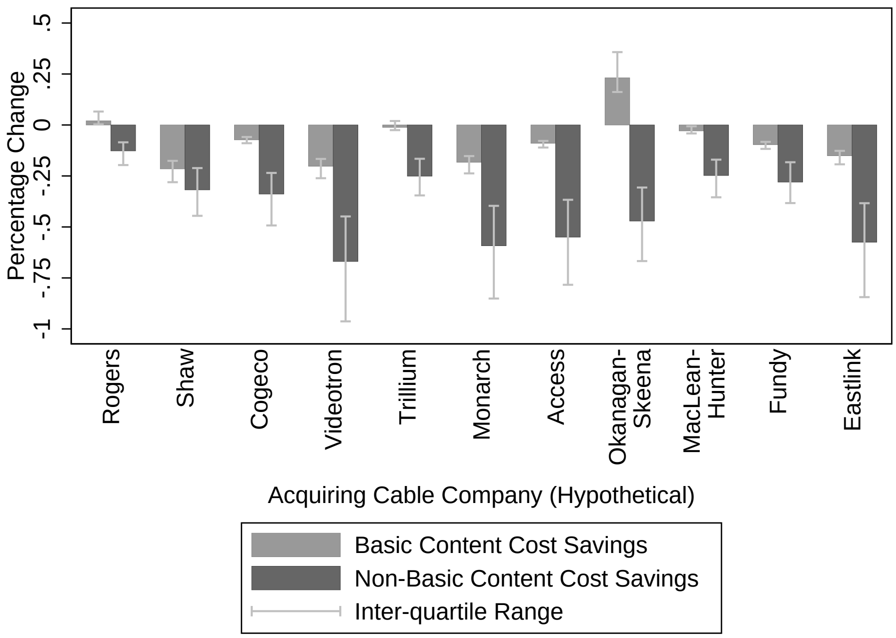


Figure 3: Predicted Cost Efficiencies from Acquisitions of Single-System Operators by Multi-System Operators

Panel A: Percent Change in Content Costs from MSO Acquisition



Panel B: Percent Change in Non-Content Costs from MSO Acquisition

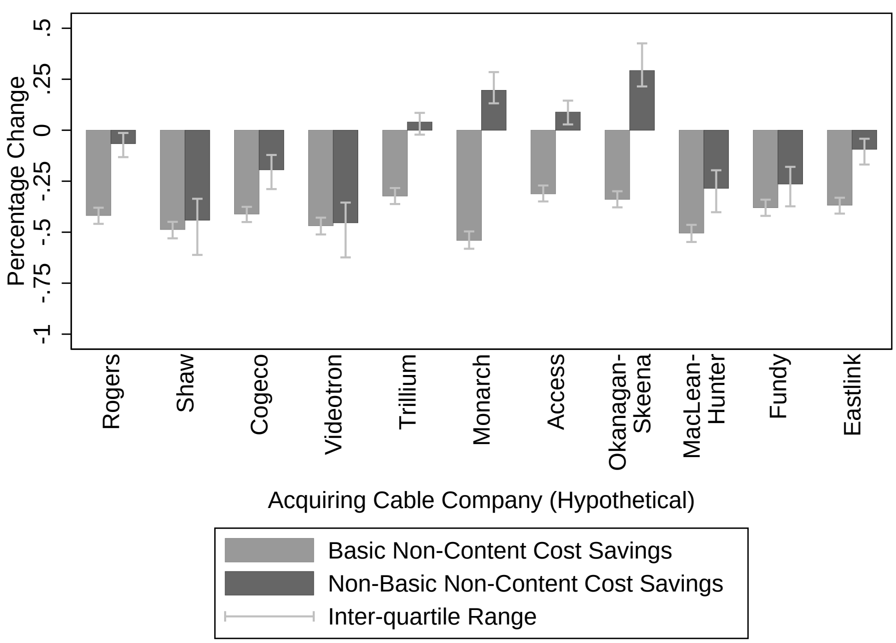


Table 1: Summary Statistics for the Demand and Cost Variables

	Licenses with Basic and Non-Basic Cable			Licenses with Basic Cable only		
	All Licenses	Large Firms' Licenses	Small Firms' Licenses	All Licenses	Large Firms' Licenses	Small Firms' Licenses
Panel A: Demand variables						
Basic price	18.76	18.40	19.04	23.42	22.74	23.51
	(4.31)	(3.66)	(4.76)	(6.02)	(5.48)	(6.09)
Non-basic price	29.68	27.92	31.10			
	(8.96)	(7.72)	(9.61)			
Basic channel count	21.86	22.53	21.32	17.34	18.00	17.25
	(5.95)	(5.68)	(6.11)	(5.95)	(6.85)	(5.83)
Non-basic channel count	9.81	12.80	7.40			
	(7.49)	(8.26)	(5.78)			
Basic market share	0.43	0.33	0.50	0.83	0.76	0.84
	(0.30)	(0.29)	(0.29)	(0.16)	(0.20)	(0.15)
Non-basic market share	0.42	0.49	0.36			
	(0.28)	(0.27)	(0.27)			
Homes passed	13352	19890	8071	1168	2025	1058
	(25103)	(30740)	(17729)	(2128)	(2761)	(2009)
Panel B: Cost variables						
Basic labor cost	3.74	3.15	4.21	4.56	3.57	4.69
	(1.91)	(1.06)	(2.27)	(3.35)	(2.35)	(3.44)
Non-basic labor cost	4.79	3.73	5.64			
	(3.63)	(1.76)	(4.44)			
Basic sales and admin cost	4.89	3.70	5.85	7.45	4.84	7.79
	(3.34)	(1.27)	(4.10)	(6.52)	(3.77)	(6.72)
Non-basic sales and admin cost	6.62	5.01	7.91			
	(5.69)	(2.39)	(7.08)			
Basic content cost	6.36	5.63	6.95	10.11	9.43	10.20
	(2.41)	(1.81)	(2.66)	(4.04)	(3.02)	(4.14)
Non-basic content cost	14.01	11.91	15.70			
	(8.10)	(6.12)	(9.06)			
Basic content cost per channel	0.32	0.27	0.36	0.64	0.61	0.65
	(0.16)	(0.12)	(0.18)	(0.32)	(0.29)	(0.33)
Non-Basic content cost per channel	2.30	1.50	3.53			
	(4.00)	(2.61)	(4.65)			
Panel C: Profits						
Basic profits	3.77	5.92	2.04	1.30	4.89	0.84
	(5.15)	(3.77)	(5.46)	(8.44)	(5.49)	(8.65)
Non-basic profits	4.26	7.26	1.85			
	(11.38)	(5.95)	(13.88)			
Basic profit margins	0.19	0.31	0.09	0.05	0.21	0.03
	(0.25)	(0.16)	(0.26)	(0.38)	(0.24)	(0.38)
Non-basic profit margins	0.15	0.27	0.06			
	(0.31)	(0.19)	(0.36)			
Panel D: Regulatory class counts						
Class 1 licenses	380	245	135	23	8	15
Class 2 licenses	607	271	336	79	17	62
Part 3 licenses	1356	530	826	657	61	596
Total number of licenses	2343	1046	1297	759	86	673

NOTES: Sample means and standard deviations (in parentheses) are presented in Panels A, B, C. All cost and profit variables are in monthly per subscriber terms, except for basic/non-basic content cost per channel, which is the monthly per-subscriber, per-channel terms. Large firms consists of the ten largest firms by national market share in 1996. Small firms are those that are not classified as large firms. Class 1 licenses have more than 6,000 basic subscribers, Class 2 licenses have between 2,000 and 6,000 subscribers, and Part 3 licenses have less than 2000 subscribers. All dollar amounts are in 1992 constant dollars.

Table 2: National Market Share and Licenses owned by Dominant Firms

	National Market Shares				Number of Licenses			
	1986	1991	1996	2001	1986	1991	1996	2001
Rogers	0.203	0.231	0.318	0.292	13	18	60	83
Shaw	0.052	0.077	0.179	0.256	16	40	92	144
Vidéotron	0.132	0.143	0.147	0.180	30	47	47	62
Cogeco	0.010	0.057	0.058	0.115	1	70	72	122
MacLean-Hunter	0.062	0.089	0	0	15	35	0	0
Videon	0.023	0.019	0.019	0	2	2	10	0
Persona	0.002	0.014	0.017	0.036	4	98	118	230
Eastlink	0.003	0.004	0.017	0.022	9	13	39	61
CF Cable	0.030	0.030	0.052	0	2	2	41	0
Fundy	0.010	0.012	0.024	0	8	13	33	0
Totals	0.449	0.598	0.701	0.842	75	210	271	411

NOTES: National market share is the total number of basic cable subscribers served by a cable company across all of its licenses divided by the total number of basic cable subscribers in the country. Cells with 0 entries are non-active or acquired cable companies. The total number of licenses in the country is 1256.

Table 3: Demand Parameter Estimates

Variable	Estimate	Standard Error
Price (α)	0.524	(0.303)
Basic cable quality (β_1)		
Constant	1.000	(0.305)
Multi-system operator	0.368	(0.112)
Rogers	-0.366	(0.115)
Shaw	0.349	(0.108)
Vidéotron	0.028	(0.009)
Cogeco	0.898	(0.274)
Non-basic cable quality (β_2)		
Constant	1.186	(0.365)
Multi-sytem operator	0.302	(0.091)
Rogers	1.001	(0.307)
Shaw	0.809	(0.248)
Cogeco	0.362	(0.111)
Vidéotron	-0.568	(0.176)
Scale of cable quality taste distribution (ψ_1)		
Constant	0.741	(0.303)
Average household income	-0.096	(0.003)
Urban density	-1.599	(0.033)
Market with basic and non-basic cable	-0.058	(0.001)
Shape of cable quality taste distribution (ψ_2)		
Constant	3.969	(0.150)
Income dispersion across households	0.838	(0.057)
Population	0.842	(0.023)
Market with basic and non-basic cable	0.552	(0.020)
Observations	5,446	
Number of licenses	3102	
GMM objective	0.006	
Degrees of freedom	36	
χ^2 critical value (5%)	50.998	

NOTES: All specifications include fixed effects for years, provinces, and individual multi-system operators. Efficient 2-step GMM robust standard errors are reported in parentheses. The scale of the distribution of the cable quality preference shocks is inversely proportional to $\rho_{mt} = \exp(y^1_{mt}\psi_1)$. The shape parameter is computed as $\kappa_{mt} = 0.1 + 14.9 \cdot \left(\frac{\exp(y^2_{mt}\psi_2)}{1+\exp(y^2_{mt}\psi_2)}\right)$, which constrains κ_{mt} to the range 0.1 to 15.

Table 4: Estimated Willingness to Pay for Basic and Non-Basic Channels

	WTP for each Basic Channel		WTP for each Non-Basic Channel	
	Mean	Std. Dev.	Mean	Std. Dev.
Single system operator	0.48	0.09	0.50	0.06
Multi system operator	0.58	0.08	0.79	0.09
Rogers	0.29	0.03	1.02	0.11
Shaw	0.61	0.07	1.08	0.12
Cogeco	0.61	0.07	1.21	0.14
Vidéotron	0.98	0.12	0.77	0.09

NOTES: Consumer i's WTP for an extra channel in bundle j for MSO k can be computed from the utility function as $\text{wtp}_{ijkmt}=\frac{t_i\times(z_{mt}\beta_{jk})}{\alpha}$. For each license-year, I simulated 1000 wtp_{ijkmt} values using the estimated demand model. For a given license-year and MSO, the mean and standard deviation of WTP can be computed with these values. The table reports the sample average of these MSO-specific WTP means and standard deviations across all license-years. All dollar amounts are in terms of 1992 constant dollars.

Table 5: Summary Statistics for Marginal Cost Estimates from Supply-side Model and Cost Data

	Licenses with Basic and Non-Basic Cable		Licenses with Basic Cable only	
	Basic Cable	Non-Basic Cable	All Markets	Unregulated Markets
Panel A: Marginal cost estimates based on cost data ($\hat{c}_{jmt}$)				
Short-run average variable cost	13.99	22.28	19.98	20.12
	[13.44,14.53]	[21.41,23.47]	[18.44,21.11]	[18.52,21.32]
Marginal cost from short-run translog cost function	9.20	18.53	18.00	17.77
	[8.37,10.15]	[17.80,19.30]	[16.90,19.44]	[16.68,19.19]
Marginal cost from long-run translog cost function	13.03	29.25	25.65	25.58
	[12.08,14.55]	[28.30,30.30]	[24.43,27.19]	[23.87,27.24]
Panel B: Marginal cost estimates inferred from the supply-side model (c_{jmt})				
Supply-side residual (ignoring basic price cap effects)	-1.75	3.69	4.88	0.05
	[-4.80,-0.41]	[0.60,5.67]	[-6.20,10.05]	[-14.47,7.41]
Supply-side residual (net basic price cap effects)	3.42	8.12	5.42	0.05
	[0.91,5.14]	[7.24,10.99]	[4.93,5.44]	[-14.47,7.41]
Licenses with supply-side residual > 0	1059	1373	410	308
Licenses with supply-side residual (net basic price cap effects) > 0	1338	1538	412	308
Total number of licenses	2344	2344	759	613

NOTES: Sample medians of the marginal cost estimates across license-years are reported. 95% confidence intervals that account for within-license dependence in costs are reported in square brackets. Confidence intervals for the inferred costs also account for variance in these costs associated with variance in the first-step demand parameter estimates. See online Appendix C for details on the bootstrapping procedures used in computing these confidence intervals. All dollar amounts are in terms of 1992 constant dollars.

Table 6: Summary Statistics for Estimates of the Content and Non-Content Components of Marginal costs from the Supply-side Model and Cost Data

	Licenses with Basic and Non-Basic Cable		Licenses with Basic Cable only	
	Basic Cable	Non-Basic Cable	All Markets	Unregulated Markets
Panel A: Marginal cost estimates based on cost data				
Content costs ($\hat{c}^x$)	0.28	0.86	0.61	0.62
	[0.26, 0.29]	[0.78, 0.95]	[0.57, 0.65]	[0.57, 0.67]
Non-content costs ($\hat{c}^{nx}$)	7.75	1.05	9.88	10.10
	[7.42, 8.03]	[0.93, 1.21]	[8.77, 10.75]	[8.65, 11.31]
Panel B: Marginal cost estimates inferred from the supply-side model				
Cost components ignoring basic price cap effects				
Content Costs (c^x)	0.89	1.75	0.63	0.61
	[0.86, 0.93]	[1.69, 1.80]	[0.61, 0.71]	[0.58, 0.70]
Non-content costs (c^{nx})	-23.46	-7.15	-7.16	-10.68
	[-27.96, -21.47]	[-11.50, -0.58]	[-17.55, -2.20]	[-21.85, -2.81]
Cost components net basic price cap effects				
Content Costs (c^x)	0.71	1.77	0.60	0.61
	[0.66, 0.81]	[1.26, 1.81]	[0.59, 0.62]	[0.58, 0.71]
Non-content costs (c^{nx})	-15.61	-7.70	-5.12	-10.68
	[-18.46, -13.52]	[-10.30, 3.61]	[-5.83, -4.75]	[-21.85, -2.81]

NOTES: Sample medians of the marginal cost estimates across license-years are reported. 95% confidence intervals that account for within-license dependence in costs are reported in square brackets. Confidence intervals for the inferred costs also account for variance in these costs associated with variance in the first-step demand parameter estimates. See online Appendix C for details on the bootstrapping procedures used in computing these confidence intervals. All dollar amounts are in terms of 1992 constant dollars.

Table 7: Marginal Cost Function Parameter Estimates

	Panel A: Estimates Using Cost Data				*Panel B: Estimates Using Inferred Costs from the Supply-side Model*			
	Content Costs		Non-Content Costs		Content Costs		Non-Content Costs	
	Basic	Non-Basic	Basic	Non-Basic	Basic	Non-Basic	Basic	Non-Basic
Firm size	-0.010^{+}	-0.206	-0.718**	-0.192	0.012^{+}	0.012	-1.636*	-4.627**
	(0.006)	(0.155)	(0.203)	(0.369)	(0.006)	(0.011)	(0.892)	(1.198)
Multi-system operator	0.032**	0.479	-1.780**	-0.919	0.184**	0.183**	1.036	3.606
	(0.021)	(0.492)	(0.596)	(1.274)	(0.022)	(0.040)	(3.376)	(4.301)
Rogers	0.058*	0.389	1.651	0.137	-0.271**	0.480**	21.868**	12.780^{+}
	(0.030)	(0.716)	(0.948)	(1.365)	(0.036)	(0.055)	(4.509)	(6.093)
Shaw	-0.071*	-0.517	0.071	-1.357	0.125**	0.442**	7.269**	-1.093
	(0.028)	(0.697)	(0.858)	(1.289)	(0.030)	(0.043)	(3.912)	(5.291)
Cogeco	-0.008	-0.771	0.547	-0.614	0.017	0.344**	16.503**	18.275**
	(0.030)	(0.868)	(0.825)	(1.368)	(0.031)	(0.059)	(4.476)	(6.250)
Vidéotron	-0.060	-1.855	0.638	-1.333	0.558**	-0.488**	-12.844	-10.082
	(0.037)	(1.023)	(1.174)	(1.673)	(0.054)	(0.053)	(7.265)	(8.422)
Average Income	0.078*	0.024	1.052	1.738	-0.919**	-2.473**	33.660**	78.925**
	(0.035)	(0.688)	(1.339)	(1.601)	(0.053)	(0.115)	(5.247)	(8.165)
Urban Density	-0.019**	-0.104	-0.277	0.079	-0.043**	-0.117**	1.016**	3.190**
	(0.005)	(0.097)	(0.157)	(0.205)	(0.007)	(0.117)	(0.928)	(1.052)
Cable system size	-0.074**	-1.080^{+}	1.316^{+}	-1.700^{+}	0.019	0.075*	-1.650	-17.685**
	(0.020)	(0.441)	(0.630)	(0.725)	(0.024)	(0.029)	(3.052)	(4.045)
Channel capacity	-0.017	0.048	0.188	-0.767*	0.004	-0.036^{+}	1.371	1.692
	(0.014)	(0.237)	(0.415)	(0.749)	(0.010)	(0.020)	(1.922)	(2.364)
Total kilometers of cable	-0.014	-0.651*	-0.159	1.020**	0.003	0.038^{+}	-0.860	-5.319*
	(0.015)	(0.248)	(0.405)	(0.756)	(0.011)	(0.020)	(1.679)	(2.286)
License with basic cable only	0.184**		0.924^{+}		-0.269**		20.296**	
	(0.022)		(0.633)		(0.032)		(2.587)	
Constant	0.268**	1.855	8.191**	0.992	1.614**	3.924**	-65.267**	-126.546**
	(0.062)	(1.233)	(1.626)	(2.634)	(0.063)	(0.133)	(9.421)	(12.207)
R-Squared	0.490	0.304	0.232	0.065	0.777	0.877	0.225	0.293
Observations	3103	2343	3103	2343	3103	2343	3103	2343

NOTES: Cluster bootstrap standard errors that account for within-license dependence are reported in parentheses. The standard errors in Panel B also account for variance in the inferred content/non-content costs associated with variance in the first-step demand parameter estimates. See online Appendix C for details on the bootstrapping procedures used in computing these confidence intervals. All specifications include fixed effects for years, provinces, and individual multi-system operators. Average income, urban density, cable system size, channel capacity and total kilometers of cable are scaled by their respective sample means and are in terms of natural logs. **, *, + indicate statistical significance at the 1%, 5%, and 10% levels.

Table 8: Reduced-form Conduct Function Parameter Estimates

	Basic Cable	Non-Basic Cable
Class 1 license	-12.493**	-13.658**
	(4.063)	(4.300)
Class 2 license	-6.541*	-5.884*
	(2.698)	(2.894)
Firm Size	-0.618	-2.440*
	(0.920)	(1.174)
Multi-system operator	6.013*	8.000+
	(3.513)	(4.847)
Rogers	10.984*	16.280**
	(4.458)	(5.751)
Shaw	10.109**	16.524**
	(3.845)	(4.933)
Cogeco	17.755**	23.177**
	(4.482)	(5.958)
Vidéotron	2.106	6.427
	(7.117)	(8.012)
⋮	⋮	⋮
Average Income	10.582*	21.576*
	(5.264)	(7.357)
Urban Density	0.672	1.389
	(0.725)	(1.077)
Channel Capacity	1.655	2.929
	(3.304)	(4.338)
System Size	3.854+	5.686**
	(1.915)	(2.480)
Total kilometers of cable	0.186	-2.607
	(1.720)	(2.366)
License with basic cable only	12.452**	
	(2.498)	
Constant	-40.575**	-66.046**
	(9.373)	(12.060)
R-Squared	0.134	0.127
Observations	3103	2343

NOTES: Cluster bootstrap standard errors that account for within-license dependence are reported in parentheses. The standard errors also account for variance in the conduct deviations associated with variance in the first-step demand parameter estimates. See online Appendix C for details on the bootstrapping procedures used in computing these confidence intervals. All specifications include fixed effects for years, provinces, and individual multi-system operators. Average income, urban density, cable system size, channel capacity and total kilometers of cable are scaled by their respective sample means and are in terms of natural logs. **, *, + indicate statistical significance at the 1%, 5%, and 10% levels.

Table 9: Impact of consolidation on cable prices, bundles, and costs

	Difference-in-Difference Estimate	Standard Merger Simulation	Modified Merger Simulation
Panel A: Cable prices and channels			
Basic price	-0.87**	-1.09	0.05
	(0.27)	[-2.08, 3.23]	[-0.79, 2.48]
Non-basic price	0.53	0.35	-6.40
	(0.99)	[-1.84, 2.16]	[-9.24, -5.10]
Basic channel count	-0.45	1.05	-4.58
	(0.37)	[-2.53, 3.47]	[-5.06, -4.23]
Non-basic channel count	1.79**	2.68	-5.28
	(0.60)	[1.66, 3.98]	[-6.94, -4.27]
Panel B: Marginal costs			
Basic content cost	-0.04*	0.09	-0.04
	(0.02)	[0.07, 0.11]	[-0.06, -0.02]
Non-basic content cost	0.42	0.31	-0.71
	(0.30)	[0.27, 0.34]	[-1.10, -0.40]
Basic non-content cost	-2.19**	-2.53	-1.88
	(0.50)	[-5.96, 0.77]	[-2.31, -1.47]
Non-basic non-content cost	-1.63**	-13.21	-2.20
	(0.58)	[-17.15, -9,23]	[-2.81, -1.64]

NOTES: Clustered standard errors at the license-level for the reduced-form merger effect estimates are reported in round brackets. 95% bootstrap confidence intervals for the predicted merger effects from merger simulations are reported in square brackets. See online Appendix C for details on the bootstrapping procedures used in computing these confidence intervals. **, *, + indicate statistical significance at the 1%, 5%, and 10% levels. All dollar amounts are in 1992 constant dollars.

Table 10: Impact of cable consolidation on monthly per-subscriber consumer surplus and profits

	Consumer Surplus	Profits
Total changes in welfare	2.39	3.86
	[0.78, 5.06]	[2.10, 6.41]
Total change in welfare without firm size effects in non-content costs	0.76 [-5.88, 6.55]	0.84 [-3.23, 5.31]

NOTES: 95% bootstrap confidence intervals are reported in square brackets. See online Appendix C for details on the bootstrapping procedures used in computing these confidence intervals. The impact of cable mergers on consumer surplus is computed as the compensating variation. All dollar amounts are in 1992 constant dollars.