A Dynamic Analysis of Consolidation in the Broadcast Television Industry *

Jessica Calfee Stahl
Board of Governors of the Federal Reserve System

March 2010

Abstract

This paper estimates a dynamic oligopoly model in order to separately identify the demand-side and cost-side advantages of consolidation in the broadcast television industry. I exploit an exogenous change in regulation that led to significant industry consolidation. Using revenue and ownership data for broadcast stations over the past ten years, I estimate the effect of ownership changes on revenue. I recover costs by examining patterns in ownership changes that are left unexplained by revenue estimation. I model firms' purchasing decisions as a dynamic game, and estimate the game using a variation of the two-step estimation method recently developed by Bajari et al. (2007). This is the first paper to estimate a model of merger activity in a dynamic, strategic setting. I find that there are significant revenue and cost advantages to consolidation. Access to a larger audience enables firms to increase per-station advertising revenue, while simply owning more stations enables firms to reduce per-station operating costs and to bargain more effectively with cable providers. A firm's ability to realize these benefits is affected by its stations' network affiliations, locations and viewer characteristics.

*Email: jessica.c.stahl@frb.gov. This is a revised version of a chapter of my Ph.D. dissertation at Boston University. I would like to thank my advisor Marc Rysman for his continued guidance and support. I would also like to thank Joshua Lustig, Randall Ellis, Iain Cockburn and Jordi Jaumandreu, as well as participants at a number of seminars, for very helpful suggestions. All errors are my own. The analysis and conclusions set forth are those of the author and do not indicate concurrence by other members of the staff, by the Board of Governors, or by the Federal Reserve Banks.

1 Introduction

Deregulation of the telecommunications industry in 1996 and 1999 has led to substantial consolidation in the local broadcast television industry. While the consolidation has provoked considerable controversy, the forces that drive it are poorly understood. This paper exploits the exogenous change in regulation, and estimates a dynamic oligopoly model so as to separately identify the demand-side and supply-side advantages of consolidation. This is the first paper to estimate a model of consolidation in a dynamic, strategic framework.

The rapid consolidation that took place following deregulation suggests that there are competitive advantages to consolidation in the broadcast television industry. Yet the vast majority of consolidation involves stations in different local markets, which are not in direct competition with one another. These markets are often located hundreds of miles apart. Therefore it is not obvious how these firms achieve either market power or economies of scale. How, for instance, does owning an NBC-affiliated station in Hagerstown, Maryland helps Nexstar Broadcasting run its Fox-affiliated station in Fort Wayne, Indiana?[1]

It is possible that firms with more stations are at an advantage in negotiations with advertisers if they offer more viewers per contract. However, this only works if firms acquire portfolios of stations that are attractive to advertisers as a group. Also, it is possible that firms with more stations are at an advantage in negotations over fees from networks and cable providers. These advantages with respect to advertisers, networks and cable providers should be reflected in higher per-station revenue for larger firms. On the other hand, it is possible that firms are able to centralize operations of stations even though they are located

[1] While most people think of local stations in terms of their network affiliations, stations are usually owned by another company altogether, such as Nexstar Broadcasting. The industry is described in more detail in Section 2.

in separate markets, often many miles apart. If firms are doing this, it should be reflected in lower per-station costs for larger firms. A final possibility is that ownership of multiple stations allows firms to smooth their revenue stream and thus reduce their vulnerability to station-specific or market-specific demand shocks.

The goal of this paper is to estimate the revenue versus cost advantages of consolidation, and to identify the characteristics of a firm's portfolio of stations that enable the firm to enjoy these advantages. I consider the number of stations a firm owns and the total population coverage of its stations, the network affiliations and locations of its stations, the similarity of its stations and their viewers, the cable providers for its stations, and the extent to which the revenue streams of its stations move together. These characteristics are likely to impact the attractiveness of the firm's portfolio of stations to advertisers, affect the firm's bargaining position with networks and cable providers, affect the ability of the firm to centralize operations of its stations, and/or affect the ability of the firm to smooth its revenue stream.

Revenue and ownership data are available for nearly all broadcast stations since deregulation began. This permits direct observation of how the revenue generated by a firm's stations responds to changes in the firm's portfolio of stations. However, I do not observe cost, so I infer cost from patterns in how firms choose their portfolios. That is, I find the effects that portfolio characteristics *must have* on cost, in order for firms' choices of portfolio characterisitics to be optimal.

This method for recovering costs is complicated by the fact that a firm's portfolio decision is both dynamic and strategic. Because the buying and selling of stations involves significant sunk costs, the effects of this period's purchasing decision will likely persist through future periods. For instance, a firm must consider that the purchase of additional stations this period will increase the number of stations that it can expect to own in all future periods, so the firm must consider the effect of the purchase not only on next period's expected profits but on all future periods' expected profits. In addition, if a firm's revenue is affected by its competitors' characteristics, then a firm will consider how its own purchasing decision will affect its competitors' decisions. Suppose, for example, that a firm hesitates to acquire additional stations even though this would lower per-station costs, because doing so would prompt its competitors to do the same. I want my model to allow for this strategic behavior, so that I do not wrongly attribute it to a lack of cost savings from owning more stations.

In order to take into account both the dynamic and strategic elements of a firm's purchasing decisions, I model the decisions as a dynamic game. I estimate the dynamic game in order to recover costs. Each firm's portfolio of stations is summarized by a vector of characteristics. In my model, a firm decides each period how to adjust these characteristics of its portfolio; this acts as an approximation of the firm's true decision about which stations to buy and sell. The firm adjusts its vector of portfolio characteristics so as to maximize the present discounted value of its expected stream of profits, basing its decision on its own current vector of portfolio characteristics, as well as those of its competitors.

The broadcast television industry is well-suited for identifying strategic behavior because firms' stations are spread across independent local markets. At any given moment, different firms are present in different sets of markets, so the nature of competition that they face varies. This feature of the industry allows me to identify the effects of strategic variables on firms' payoffs, because I observe within-time variation in the values of these strategic variables.

Until recently, it was virtually impossible to estimate dynamic games such as this for computational reasons. However, several authors have recently developed methods for estimating the parameters of a dynamic game that avoid computing the equilibrium even once. I use a variation of the two-step method proposed by Bajari et al. (2007), which allows for continuous choice variables. In the first stage, I use revenue data to estimate the effects of firms' portfolio characteristics on per-station revenue. The second stage uses ownership data to estimate the strategies of firms, that is, how they adjust their portfolio characteristics each period. I find the effects that portfolio characteristics must have on per-station costs such that the observed strategies are consistent with profit maximization. The result is a set of estimates of both the revenue and cost effects of consolidation.

This paper develops an innovative way to estimate the demand-side and supply-side advantages of consolidation in the absence of cost data. As far as I know, it is the first paper to estimate a model of merger activity in a dynamic, strategic framework. My approach allows me to isolate the relationships of interest in an industry that each year witnesses hundreds of firms considering the purchase or sale of over one thousand stations.

My results suggest that the advantages to consolidation in this industry work along a number of very different dimensions. First, I find that firms are able to raise more revenue per station if all of their stations together reach a larger total population, suggesting that firms can charge higher rates if they can offer advertisers more viewers per contract. I find that firms are better able to do this if their markets are closer together and similarly sized, and if they can offer stations with a variety of network affiliations.[2] I also find evidence that a firm can extract higher fees from a cable provider if its stations represent a large

[2]Note that if advertisers want to reach a large audience, they can always advertise directly through a network. So a firm may not have much to offer an advertiser if they essentially just own, for instance, a subset of NBC's stations.

percentage of the broadcast stations carried by the cable provider.

There are also cost-side advantages to consolidation. I find that firms face lower per-station costs if their own a greater number of stations, suggesting that firms are able to combine the operations of stations that if they are located in different markets. Firms can achieve greater economies of scale if their markets are similar in terms of their size, demographics and income levels. Consolidation also enables firms to reduce their exposure to revenue fluctuations. Yet I find that firms incur significant sunk costs when they purchase stations, suggesting that dynamics are important in this industry. Also, a firm's revenue is significantly affected by the portfolio characteristics of its competitors.

The rest of the paper proceeds as follows. Section 2 describes the relevant literature, Section 3 describes the broadcast television industry and Section 5 describes the data that I use. Section 4 describes the dynamic oligopoly model and Section 6 explains my estimation strategy in detail. Section 7 explains my results and Section 8 concludes the paper.

2 Related Literature

There is a substantial theoretical literature seeking to understand why some industries initially support many competitors but then undergo significant consolidation or "shake-outs."[3] Some of these studies focus on the evolution of an industry from a highly competitive industry to an oligopoly, while others focus on the features of an industry that characterize its inherent or "natural" level of competitiveness in the long run. Because the broadcast television industry was

[3] These studies include, but are not limited to, Jovanovic (1982), Shaked and Sutton (1983), Shaked and Sutton (1987), Klepper and Grady (1990), Shaked and Sutton (1990), Hopenhayn (1992), Petrakis and Roy (1999), Klepper and Simons (2000), and Klepper (2002).

not allowed to consolidate until the late 1990s, it is impossible to know what its natural evolution would have been. The rapid consolidation that followed deregulation suggests that there are demand-side and/or supply-side features of the industry that make it naturally concentrated rather than highly competitive. This paper takes advantage of a shake-out triggered by the exogenous deregulation in order to identify those features.

Shaked and Sutton (1987) develop the idea that certain features of an industry's production technology can lead the industry to be concentrated regardless of the size of demand. Specifically, they look at the relationship between a firm's cost and the quality of its product. My results suggest that this relationship is important in the case of the broadcast television industry. Firms with greater population coverage appear to offer advertisers an inherently higher-quality product, suggesting that consolidation changes the relationship between cost and quality for broadcasting firms.

There is relatively little empirical research on mergers and consolidation despite the obvious practical importance of the topic. A persistent problem is the difficulty of identifying the factors that drive consolidation because merger activity is endogenous. Event analysis uses the exact timing of the announcement of a merger as a source of exogeneity. Also called event studies, these look at how merger announcements affect the stock market values of the firms affected by the merger, in order to understand the effects of the merger. For instance, Knapp (1990) does an event analysis of the airline industry and concludes that mergers in this industry increase market power.

Other papers make use of exogenous sources of variation in merger activity prompted by changes in regulations. For instance, Paul (2003) and Pesendorfer (2003) both estimate the cost structure of an industry before and after consolidation, making use of revisions to merger guidelines to argue that merger activity was exogenous. Paul looks at the beef-packing industry, while Pesendorfer looks at the paper industry. Both authors conclude that consolidation alters the cost structure of firms. Kim and Singal (1993) make use of a period of time when airline mergers went uncontested, and find a positive effect of airline mergers on prices.

A few papers look specifically at consolidation in the broadcast television industry following deregulation. In 2006, the Federal Communications Commission (FCC) commissioned ten studies to estimate the effect of media consolidation on "diversity, competition and localism."[4] Crawford (2007) considers the impact of consolidation on the quantity and quality of local television programming. Shiman (2007) studies the relationship between ownership structure and the quantity of news programming, and Hammond et al. (2007) study the relationship between consolidation and minority and female ownership of television stations. Results are largely inconclusive; they depend on the methods of measurement and estimation, and are often statistically or economically insignificant.

In all of these studies, ownership changes are treated as strictly exogenous, and the mechanisms through which they affect market outcomes are left largely unexplained. There is very little empirical work that endogenizes consolidation or firms' merger decisions. An exception is the recent literature that models mergers as matching games between firms. The approach is similar to mine in that it uses merger decisions to make inferences about the parameters of firms' payoff functions or to reveal firms' incentives. Akkus and Hortacsu (2007) model bank merger decisions to reveal what features of the merging banks enable them to create more value through merger. Park (2008) uses a matching model

[4] "FCC Names Economic Studies to Be Conducted As Part of Media Ownership Rules Review," November 22, 2006. See http://www.fcc.gov/ownership/actions.html.

to analyze manager incentives in mutual fund acquisitions.[5] However, matching models tend to suffer from a dimensionality problem. As far as I know, no matching model has incorporated dynamics, and even the incorporation of strategic effects has been limited. For instance, Fox (2007) suggests a way to avoid the curse of dimensionality in a matching model; the method essentially involves the assumption that a given pair of matching agents do not care about the other matches that occur in the market, so long as they prefer their own match to any other match they could have made. This assumption is reasonable in the context of, for instance, a marriage market. However, I am not willing to make this assumption, given that broadcast television firms may be directly affected by mergers between other firms in the industry. For this reason, I do not model mergers as a matching game. However, the incorporation of strategic and dynamic effects in a matching model is an interesting area for future research.

Just recently, Jeziorski (2010) has looked at the merger wave in the radio industry following the 1996 Telecom Act, in a dynamic, strategic setting. In radio, there were many within-market mergers, and Jeziorski focuses on within-market competition and consolidation. In this setting, the advantages of consolidation are obvious: market power and straightforward economies of scale. Jeziorski (2010) seeks to estimate welfare effects, taking into account post-merger format changes. In order to deal with the dimensionality problem, Jeziorski makes strict assumptions about the ordering of mergers. Also, since markets are treated as independent and across-market mergers are ignored, the number of potential mergers that must be considered in the model is greatly reduced. This approach makes sense in radio, where the bigger story is within-market consolidation, but would not make sense in the television industry, where the bigger story is

[5]See Sorenson (2007), Fox (2007) and Ho (2007) for other examples of matching game estimation, applied to settings other than mergers.

across-market consolidation.

On the methodological side, I use a variation of the method proposed by Bajari et al. (2007) to estimate my dynamic model of consolidation. Recent papers that have used this or similar methods to estimate dynamic games include Collard-Wexler (2008) which estimates the impact of demand fluctuations in the ready-mix concrete industry, Ryan (2009) which estimates the costs of environmental regulation in the Portland cement industry, Snider (2008) which estimates predatory incentives in the airline industry, and the previously mentioned Jeziorski (2010). My variation of the method proposed by Bajari et al. (2007) uses first-order conditions as the basis for estimation, as suggested by Berry and Pakes (2000).

3 Broadcast Television Industry

Originally, broadcast television stations could only be viewed by those who lived within reach of the station's over-the-air signal. Now most households in the United States have cable television, so they view their local broadcast stations through their cable provider; with very few exceptions, though, these stations are only provided locally. The FCC's "must-carry" laws force cable providers to carry local stations, but broadcast stations can opt out of the "must-carry" provision and ask cable providers to pay a fee in exchange for the right to carry the station.

Historically, broadcast stations were locally owned, but as major national networks emerged, the government became concerned that local programming would be dominated by national interests. The FCC put into place a number of restrictions that were meant to encourage local ownership; these are discussed in detail later. Stations were thought to be more able to respond to the needs of the local community if they were not owned by or beholden to national networks. For example, the FCC's 1941 "Report on Chain Broadcasting" maintains that "[a] station licensee must retain sufficient freedom of action to supply the program and advertising needs of the local community. Local program service is a vital part of community life."

The total number of stations in the country is effectively fixed over time because licenses for new stations are rarely issued, and stations virtually never give up their licenses. There are 1,251 full-power commercial stations in the country, across 210 markets.[6] Many stations maintain an affiliation with a national network; the station agrees to broadcast the network's prime-time programming and advertising, usually in exchange for a fee. By maintaining affiliated stations in most markets across the country, the major networks have national coverage without violating FCC ownership rules. Four networks have essentially national coverage: ABC, CBS, Fox and NBC. The WB and UPN merged in 2006 to form CW, which has emerged as a competitor to the major four. There are also a handful of smaller networks, the most significant of which are Ion and the Spanish-language networks Univision, Telemundo and Telefutura.

The Telecommunications Act of 1996 was the first major overhaul of telecommunications policy since the Communications Act of 1934 created the FCC nearly 62 years earlier. The 1996 Act came in the wake of the breakup of AT&T, when there seemed to be some consensus among policy-makers that the telecommunications industry was over-regulated. Broadcasting companies argued that they were struggling to compete with cable television and home video, and ex-

[6]Nielsen Media Research divides the country into 210 Designated Market Areas (DMAs) that are essentially mutually exclusive. A typical market has between five and eight stations; Boston and Houston have the most with 14 each, whereas a number of small towns have just one station each.

erted pressure on the FCC to loosen restrictions on ownership of stations. The 1996 Act removed the limit on the total number of stations that a single entity could own (previously 12) and increased from 25% to 35% the portion of the national population that a single owner's stations could cover. Further deregulation in 1999 allowed joint ownership of two stations with overlapping coverage[7] (previously disallowed) as long as they were not in the same market, and allowed joint ownership of two stations in the same market (previously disallowed) as long as that left eight independent full-powered commercial stations in the market and neither of these two stations were ranked in the top four in the market in terms of viewing share.[8] In 2003 the FCC sought to increase the coverage cap from 35% to 45%, but Congress intervened and chose a cap of 39%.[9]

The industry underwent significant consolidation in the ten years following deregulation. Views on the effects of this consolidation differ dramatically. For instance, the Common Cause Education Fund argues "that profit-driven media conglomerates are investing less in news and information, and that local news in particular is failing to provide viewers with the information they need in order to participate in the democracy."[10] On the other hand, the National Association of Broadcasters (NAB) maintains that "[t]he real threat today to locally-oriented services, including costly services such as local news, is not the joint ownership of broadcast stations..., but the stations' continuing challenge to maintain their

[7]Technically, their "Grade B contours" are now allowed to overlap. "Grade B" is a measurement of the strength of a television signal as it arrives at a viewer's antenna. A station's "Grade B contour" is the geographic area in which it is predicted that a consumer with an outdoor rooftop receiving antenna can pick up a signal of Grade B intensity.

[8]This last rule is referred to as a "duopoly rule," but obviously does not refer to a duopoly as economists define one.

[9]CBS and Fox already had 39% coverage and had received waivers from the FCC.

[10]"The Fallout from the Telecommunications Act of 1996: Unintended Consequences and Lessons Learned" by the Common Cause Education Fund, May 9, 2005.

economic vibrancy in the face of multichannel and other competitors that are not constrained by restrictions on local ownership structure."[11] Broadcasting companies argue that ownership deregulation was necessary for the survival of broadcast television and that further deregulation is in order. Congress has explicitly allowed for this possibility; it directed the FCC to re-evaluate remaining regulations every two years to "determine whether regulation is no longer necessary in the public interest."[12] Yet the debate rarely addresses the question of *how* station owners are helped by consolidation; in fact, this is not well understood.

This paper estimates the revenue and cost advantages of consolidation. It is generally thought that increased revenue following consolidation has negative welfare implications, while lower cost following consolidation has positive welfare implications. However, the usual intuition does not necessarily hold in this industry. Because remaining FCC regulation prevents individual firms from dominating local markets, increased revenue following consolidation is most likely not evidence of market power. If bigger firms attract more advertising revenue per station, it is most likely because they offer advertisers more viewers per transaction. The advertiser's outside option has not changed; the bigger firm is simply offering the advertiser a more valuable product. Alternatively, if bigger firms extract higher fees from networks and/or cable providers, this is simply a transfer of surplus to the broadcasting firm from another party.

On the other hand, it is often argued that local television programming is a public good. Thus a reduction in the costs incurred in the production of local programming may reveal under-provision of a public good. For these reasons,

[11] Comments of the National Association of Broadcasters Before the Federal Communications Commission, October 22, 2007, p.3.

[12] Section 11(2)(a) of the Telecommunications Act of 1996. This was later changed to every four years.

I hesitate to use my results to draw strong conclusions about welfare. However, as discussed in Section 2 on related literature, studies of this industry have attempted to directly measure the reduced-form effects of consolidation on particular outcomes so as to draw welfare conclusions, and have been inconclusive. It is difficult to understand the mechanisms through which consolidation affects welfare without understanding why consolidation has taken place. This paper attempts to fill that void, by identifying the demand-side and supply-side factors that have driven consolidation in this industry.

4 Model

I estimate a model in which a firm chooses each period how to adjust the characteristics of its portfolio of stations. The firm's static per-station revenue and per-station cost are functions of its own portfolio characteristics as well as those of its competitors. I assume that firm i's payoffs depend on its competitors' portfolio characteristics only through the average values of these characteristics among all of its direct competitors, where a direct competitor is defined as a firm that owns a station in one of firm i's markets. A firm incurs sunk costs when it adjusts its portfolio, so this period's adjustments affect payoffs and decisions in future periods. Firms choose their portfolio each period so as to maximize the present discounted value of their expected stream of profits as a function of their own current portfolio characteristics and their competitors' current average portfolio characteristics. This framework takes into account the strategic and dynamic elements of the firm's decision. By summarizing the firm's portfolio by a vector of characteristics, the model loses the ability to predict individual transactions. However, the advantage of this approach is that I can flexibly estimate the effects of portfolio characteristics on revenue and cost in a game consisting of hundreds of firms facing thousands of choices each period.

4.1 Revenue and Cost

Let firms be indexed by $i = 1, \ldots, I$ and years be indexed by $t = 1, \ldots, \infty$. The vector of characteristics describing firm i's portfolio in period t is denoted $char_{it}$, and the vector of averages of those characteristics among firm i's competitors in period t is denoted $char_{-it}$. Firm i's static per-station revenue in period t is:

$$r_{it} = \gamma(char_{it}, char_{-it}; r) + \mu_i^r + \nu_{it} \tag{1}$$

where $\gamma(\cdot)$ is a reduced-form function of $char_{it}$ and $char_{-it}$, parameterized by the vector of revenue coefficients r; μ_i^r is a firm-specific revenue constant (a firm fixed-effect) and ν_{it} is a demand shock distributed $N(0, \sigma_i^r)$.

Firm i's static per-station cost in period t is a function of its own portfolio characteristics, subject to adjustment costs[13]:

$$c_{it} = \zeta(char_{it}, \Delta n_{it}, \omega_{adj_{it}}; c) + \mu_i^c + \omega_{per_{it}} \tag{2}$$

where $\zeta(\cdot)$ is a reduced-form function of $char_{it}$ and adjustments to a firm's number of stations Δn_{it}, parameterized by the vector of cost coefficients c; μ_i^c is a firm-specific cost constant, $\omega_{per_{it}}$ is a shock to per-station cost distributed $N(0, \sigma^{c_{per}})$, and $\omega_{adj_{it}}$ is a shock to adjustment cost distributed $N(0, \sigma^{c_{adj}})$.

The firm's total revenues and costs are simply:

$$R_{it} = n_{it} * r_{it}$$

[13]Adjustment costs are the sunk costs incurred when buying or selling stations, such as transaction costs. The adjustment cost does not include the price paid for a station, because this is not necessarily a sunk cost. A firm's choice to have its capital invested in a station is part of the annual opportunity cost of holding the station; this will be included in the per-station cost constant. This is an ongoing cost, not a sunk cost incurred when the purchase takes place.

$$C_{it} = n_{it} * c_{it}$$

where n_{it} is simply firm i's number of stations. The firm's static profit is:

$$\pi_{it} = R_{it} - C_{it} = n_{it}(r_{it} - c_{it})$$

4.2 Strategies and Equilibrium

Since a firm's static profit depends on its competitors' portfolios only through $char_{-it}$, firm i's state in period t can be fully described by the state vector S_{it}:

$$S_{it} = \{char_{it}, char_{-it}, \Delta n_{it}, \mu_i^r, \mu_i^c, \nu_{it}, \omega_{per_{it}}, \omega_{adj_{it}}\}$$

A firm's strategy maps from its state vector this period to the vector of its own portfolio characteristics next period:

$$\sigma_i : S_{it} \rightarrow char_{it+1}$$

The firm chooses σ_i to maximize the present discounted value of its expected stream of profits given the strategies of its competitors. This value can be written:

$$V(S_{it}|\sigma_i, \sigma_{-i}) = E\left(\sum_{\tau=0}^{\infty} \beta^{\tau} \pi_{it+\tau} | S_{it}, \sigma_i, \sigma_{-i}\right) \tag{3}$$

The equilibrium concept is Markov-perfect Nash equilibrium; a firm's strategy depends only on its current state vector. The equilibrium can thus be defined as the strategy σ^* that maximizes the value of state S:

$$\sigma_i^* = argmax_{\sigma_i} V(S_{it}|\sigma_i, \sigma_{-i}) \tag{4}$$

In the data, we see that some firms grow rapidly while others stay very small or exit. In my model, heterogeneity in growth patterns across firms comes from three sources. First, some firms are inherently more profitable than others; they have higher μ_i^r or lower μ_i^c. The return to growth is higher for these firms because they generate higher per-station profit; therefore they are more likely to grow. Second, firm-specific shocks to adjustment costs $w_{adj_{it}}$ generate heterogeneity in optimal growth. Lastly, portfolio characteristics affect per-station profitability through $r(\cdot)$ and $c(\cdot)$; thus a firm's portfolio in period t will affect its optimal size in future periods. So, in the model, firms that are more efficient and/or receive favorable adjustment cost shocks will grow, and then once big, will be more likely to keep growing. This story fits the growth patterns seen in the data.

5 Data and Specification

I acquired a proprietary dataset from BIA Financial Network Inc (BIAfn), a telecommunications research firm. I observe each station's market, owner, revenue, population coverage, and network affiliation annually from 1996 through 2007. For each market, I observe the average per capita income and the demographics of the population. Each station is uniquely identified by its call letters. Some effort was required to standardize owner names and create a unique identifier for each firm. The dataset as provided by BIAfn is at the level of the station; I use this to construct a panel dataset at the level of the firm. I also collect data on the cable providers that are present (and have at least 1,000 subscribers) in each market in the country in 2009; unfortunately I only have a snapshot, rather than time series data on cable coverage.

A look at the data confirms that the broadcast television industry underwent significant consolidation in the ten years following deregulation. For the purposes of identification, I would like to see the industry adjusting from a pre-deregulation equilibrium to a post-deregulation equilibrium. Fortunately, the data do suggest that the industry was in equilibrium in 1995 (before the change in regulation) and that by 2005 or so had reached a new equilibrium. Immediately following deregulation, stations began changing hands at a much greater rate than they had been. Figure 1 shows the number of transactions (sales of stations) on a yearly basis before and after deregulation. Purchasing activity picked up dramatically in 1996 (the dashed line), then gradually declined over the sample period. By 2005, station sales had fallen back to pre-deregulation levels and the market structure in the industry had stabilized.[14] I am using this period of transition from one equilibrium to another to estimate the effects of consolidation. The hope is that, prior to 1996, firms would have wanted to make these station purchases/sales, but they were constrained by regulations. Once (some) constraints were lifted, it took a number of years for the industry to reach a new equilibrium. In estimation, I allow for adjustment costs to explain this.

In 1995, just before deregulation, most firms held only one station and only three firms held more than ten stations. After deregulation, firms that already held a relatively large number of stations tended to expand their portfolios. For instance, of the eight firms that owned ten or more stations in 1995, all but three of them owned more than twenty stations by 2007. On the other hand, single-station owners tended to exit the industry. Nearly 70 percent of the firms that owned only one station in 1995 had sold their station by 2007; only one owned more than three stations by 2007. By 2007, 25 firms owned more than ten stations each, and the majority of stations were in the hands of these large owners. The greatest number of stations held by one firm in 2007 was 60. Figure 2 and Figure 3 show how the ownership structure of the industry

[14] That is, the market structure in 2007 looked very much the same as it did in 2005. Small owners had mostly stopped selling and large firms had mostly stopped buying stations by 2005.

changed over these years. The trend toward consolidation suggests that owning more stations puts firms at a competitive advantage.

How are owners of multiple stations at a competitive advantage? The goal of estimation is to find the features of a firm's portfolio that enable it to lower its per-station costs or raise its per-station revenue. These are features that describe the firm's stations *as a group*, rather than characterizing the firm's stations themselves, because I am trying to capture how joint ownership of these stations affects their profitability.[15]

Broadcast stations earn most of their revenue from advertising. An advertiser wanting to reach a large number of viewers may prefer to negotiate a single contract rather than negotiate many separate contracts with individual stations. Therefore a firm that reaches a large number of viewers may be able to charge higher advertising rates per station if their stations are attractive to advertisers as a group. In estimation, I include in the revenue function any portfolio characteristics that are likely to affect the attractiveness of the portfolio to advertisers. Advertisers primarily care about number of viewers, so a firm's total population coverage is included in the revenue equation. Advertisers may also care about the number of different stations reached, so this is included as well.[16] Advertisers may also care about the heterogeneity of the firm's stations (and therefore viewers), so I include various measures of this: average distance between sta-

[15]Factors that affect individual station profitability are hopefully captured by station/portfolio fixed effects.

[16]Advertisers presumably also care about the number of different markets reached. However, since firms are quite restricted in their ability to own multiple stations per market, a firm's number of markets is usually the same as its number of stations. There are a few exceptions, but I cannot distinguish between the effects of the two variables in regressions because they are so highly correlated (correlation = 0.99).

tions, regional concentration of stations,[17] network concentration of stations,[18] variance of stations' coverage areas,[19] demographic heterogeneity,[20] and income heterogeneity.[21]

Stations can also earn revenue through fees from networks and cable providers. When a station is affiliated with a national network, it agrees to broadcast the network's prime-time programming (and the network's pre-arranged advertising) in exchange for a fee. If a firm has more bargaining power with a network, it may be able to negotiate higher fees. As a measure of a firm's potential bargaining power with networks, I construct a measure a firm's level of "network domination:" I calculate the percentage of each network's stations that the firm owns (weighted by revenue), and sum this across all networks. Similarly, I construct

[17]This is the Herfindahl index as applied to the firm's concentration among regions of the country, divided by the expected value of this index given the number of stations held by the firm. The regions are New England (ME, NH, VT, MA, RI, CT), the Mid Atlantic (NY, NJ, eastern PA), the Midwest (western PA, WI, MI, IL, IN, OH, ND, SD, NE, KS, MN, IA, MO), the Southern Atlantic (DE, MD, VA, DC, WV, NC, SC, GA, FL), South Central (KY, TN, MS, AL, OK, TX, AR, LA), Mountain (ID, MT, WY, NV, UT, CO, AZ, NM) and Pacific (WA, OR, CA, AK, HI).

[18]This is the Herfindahl index as applied to the firm's concentration among the national network affiliations, where the networks are those name in Section 3. This is then normalized by the expected value of the index given the number of stations held by the firm.

[19]This is calculated as the coefficient of variation in population coverage across the stations in the portfolio.

[20]This is measured as the average Euclidean distance between the demographic vectors of a firms' stations. A demographic vector is defined as: {percent of households that are Caucasian; percent African-American; percent Asian; percent Hispanic}. So, if a firm has one station in Chicago, where the demographic vector is $\{64, 18, 5, 13\}$, and Houston, where the demographic vector is $\{56, 12, 4, 28\}$, then the demographic heterogeneity variable is $[(64-56)^2 + (18-12)^2 + (5-4)^2 + (13-28)^2]^{\frac{1}{2}} = 18$.

[21]This is calculated as the coefficient of variation in the average income per household across the stations' populations.

a measure of a firm's level of "cable domination" in order to capture bargaining power in negotiations with cable providers. I include quadratic terms of all regressors, to allow for concavity (or convexity) in the effects of these portfolio characteristics on per-station revenue. Also included in a firm's revenue equation are the average portfolio characteristics of the firm's competitors, to allow for strategic effects.

In the cost equation, I include portfolio characteristics that might affect the firm's ability to centralize operations of its stations. These include the firm's total number of stations, the average distance between its stations and regional concentration of its stations. Firms may be better able to centralize the production and purchase of programming if their stations have similar viewers, so I also include the variance of the stations' coverage areas, as well as demographic heterogeneity and income heterogeneity. Also, there are presumably costs associated with dealing with networks and cable providers, so I include the total number different networks that the firm's stations are affiliated with, and the total number of different cable providers in the markets in which the firm has a presence.

Recall that in estimation, cost is essentially a residual, serving to explain the behavior of firms that is left unexplained by revenue. A firm's costs of operation include, for instance, the cost of raising capital, the cost of being exposed to risk, and so on. I include in the cost equation the coefficient of variation in the firm's stream of revenue. The coefficient on this regressor will reflect the extent to which firms forgo expected revenue in order to be able to smooth revenue across time, which reveals the costliness of enduring large year-to-year fluctuations in revenue. This is of interest because another possible advantage of consolidation is that larger firms may be able to better smooth revenue across time. Indeed, as shown in Figure 4, firms with larger numbers of stations face less volatility in their revenue stream.

I also allow for adjustment costs in the purchase and sale of stations. There are likely to be some sunk costs involved in buying and selling television stations: legal costs, time and money invested in negotiating a price and deciding on the sale, the cost of transferring ownership, etc. I allow for asymmetry in adjustment costs; that is, the cost of buying a station is allowed to be different than the cost of selling a station. These are likely to affect the total costs incurred by the firm, rather than the average cost of running one of the firm's stations. Since the cost equation is for per-station cost, I divide the adjustment by the number of stations to get the total firm-level cost of adjustment.

Table 1 shows the summary statistics for all of the portfolio characteristics that are included in the revenue and/or cost equations. The characteristics which show the greatest variation (relative to their mean) are number of stations, population coverage, network domination and cable domination.

Estimation involves only those firms that own more than one station. Out of a total of 1,251 stations, 131 are dropped because they do not providee revenue data. I drop those firms that own major networks because these firms are likely to pursue different strategies, and face different trade-offs and economies of scale, than other station owners.[22] Some firms were dropped because their stations did not have enough revenue data to enable me to identify firm fixed effects in the revenue regression. Ultimately I am left with a total of 244 firms, and 1,799 firm-year observations, distributed fairly evenly across the years 1996 to 2007.[23] In estimation, revenue is adjusted to constant 2008 dollars.

[22]This involves dropping seven firms, which in 2009 owned a total of 180 stations altogether (out of 1,251 in the country). Note that this does not involve dropping stations that are affiliated with major networks; just those that are *owned* by major networks.

[23]Strategic variables (average portfolio characteristics among competitors) are calculated before firms/stations are dropped.

Note that a firm's underlying portfolio decisions affect its revenue and cost through different variables. For instance, when a firm chooses the overall size of its portfolio, this affects revenue through population coverage, but affects cost through number of stations (since there is a marginal cost associated with operating an additional station, but no marginal cost associated with reaching an additional viewer) as well as through adjustment costs associated with additional stations. Also, when a firm chooses its stations' network affiliations, this affects revenue through the network domination and concentration variable, but affects cost through the total number of different network affiliations. Similarly, when a firm chooses which markets to be in, this affects revenue through cable domination and strategic effects, but affects cost through the total number of cable providers (and affects both revenue and cost through the similarity of the markets; that is, their distance apart, regional concentration, relative sizes, viewer demographics and viewer incomes). The exclusion of some revenue-side variables from the cost equation and vice-versa is crucial for the separate identification of meaningful revenue and cost estimates.

6 Estimation

6.1 Empirical Strategy

The goal of the paper is to estimate the paramaters of the revenue and cost functions. Because revenue is observed, the revenue function can be estimated by regression. Costs are not observed, so I find the cost function that satisfies the first order conditions derived from the dynamic game in which firms choose their portfolio characteristics to maximize the present discounted value of their streams of expected profits, given their competitor's choices. The use of first order conditions as moments to estimate a dynamic game is suggested by Berry and Pakes (2000). However, with continuous choice variables, estimation of the first order conditions in expectation is difficult. I apply the simulation method which Bajari et al. (2007) propose to estimate ex ante value functions; here I estimate ex ante partial derivatives of value functions.

6.2 Revenue Estimation

The first step is to estimate the revenue equation (1).[24] The equation is at the level of the firm-year. The dependent variable is the average revenue generated by firm i's stations in period t. The regressors are the characteristics of i's own portfolio that were discussed in Section 5 and the average characteristics among the portfolios of i's direct competitors. The firm-specific standard error of the demand shock σ_i^r is estimated as the standard error of the regression residual across all of firm i's observations.

One complication is that I am eliminating station identity from the model, yet stations are heterogeneous. Failing to take into account station heterogeneity might lead to biased coefficients on portfolio characteristics in the revenue equation, which would in turn lead to biased cost coefficients in the second stage of estimation. This would happen if the revenue potential of a station is correlated with the characteristics of its owner's portfolio which are included in estimation.[25] In order to deal with station heterogeneity, I include as a regressor

[24]Because the results of the revenue equation are inputs in the dynamic model used to estimate the cost equation, I am restricted in the specification of the revenue equation. Most notably, I must use a linear specification, and must keep the number of regressors within reason. This will be discussed in more detail in the section describing the estimation of the cost equation.

[25]For example, suppose that owning more stations increases revenue with no effect on cost. But also suppose that firms with more stations tend to own smaller stations; that is, stations

in the revenue equation the average revenue generated by firm i's period-t stations across all years and owners. This can be thought of as a "portfolio fixed effect:" the average revenue of i's portfolio of stations across all years, regardless of ownership. This is much like station fixed effects, but even more stringent. I also include this value in the cost equation, to allow for the (likely) possibility that stations that generate more revenue also cost more to operate.

In addition to portfolio fixed effects, I also include firm and year fixed effects. Therefore I am effectively looking at variation within the firm-station-year. Given this, the most important identification assumption is that there is no trend in a station's revenue which induces a firm with particular portfolio characteristics to purchase it (no reverse causality). Since the timing of station purchases was largely the result of an exogenous change in regulation, this is likely to be a reasonable assumption.

6.3 Cost Estimation

Estimation of the cost equation (2) is considerably more complicated. Costs are not observed, so I look at the strategies employed by firms - that is, their portfolio choices each period - in order to recover costs. I find the cost coefficients that satisfy the first order conditions and thus make the estimated strategies optimal. The firm incurs sunk costs when it buys and sells stations, so its portfolio adjustments today affect all future payoffs. Therefore the firm adjusts

that, for unobservable reasons, generate less revenue and incur less cost regardless of ownership. Then the coefficient on n_{it} in the revenue equation will be biased downward. When looking at firm strategies to infer costs, I will be searching for the cost coefficient that explains the firm's tendency to buy up stations despite the (spurious) fact that this does not improve revenue. The cost coefficient on n_{it} will be biased downward as well; it will look like owning more stations lowers costs more than it actually does.

its portfolio characteristics so as to maximize its *stream* of profits, not just its static profits today.

Let the portfolio characteristics used in estimation be indexed by j, such that $\{char_{ijt}\}_{j=1,\ldots,J} \in char_{it}$ and $\{char_{-ijt}\}_{j=1,\ldots,J} \in char_{-it}$. Recall that the firm's choice variables are its own portfolio characteristics $\{char_{ijt}\}_{j=1,\ldots,J} \in char_{it}$. The first order conditions for optimal portfolio choices are:

$$E\left[\sum_{\tau=0}^{\infty} \beta^{\tau} \frac{\partial \pi_{it+\tau}}{\partial char_{ijt}} \mid S_{it}\right] = 0 \quad \forall \quad i,j,t \tag{5}$$

Equation (5) says that the marginal effect of the portfolio characteristic $char_{ijt}$ on the present discounted value of firm i's expected stream of profits must equal zero. The choice of the portfolio characteristic today $char_{ijt}$ affects expected future profits because it affects expected future states. So, we can write the first order conditions as:

$$E\left[\sum_{\tau=0}^{\infty} \beta^{\tau} \left(\frac{\partial \pi_{it+\tau}}{\partial S_{it+\tau}}\right)' \frac{\partial S_{it+\tau}}{\partial char_{ijt}} \mid S_{it}\right] = 0 \quad \forall \quad i,j,t \tag{6}$$

The effect of current portfolio characteristics on future states ($\partial S_{it+\tau}/\partial char_{ijt}$) works through firm strategies, and recall from Section 4 that firm strategies depend only on the current state vector. So we can break down the first order conditions as follows:

$$E\left[\sum_{\tau=0}^{\infty} \beta^{\tau} \left(\frac{\partial \pi_{it+\tau}}{\partial S_{it+\tau}}\right)' \frac{\partial S_{it+\tau}}{\partial S_{it+\tau-1}} \frac{\partial S_{it+\tau-1}}{\partial S_{it+\tau-2}} \cdots \frac{\partial S_{it+2}}{\partial S_{it+1}} \frac{\partial S_{it+1}}{\partial char_{ijt}} \mid S_{it}\right] = 0 \quad \forall \quad i,j,t \tag{7}$$

6.3.1 Estimation of First Order Conditions

In this section, I discuss the estimation of the expressions in (7) for a given value of the cost parameter vector c. In the next section, I discuss the search for c^* that sets this expression equal to zero, satisfying the first order condition (7).

In the estimation of (7), the first step is to estimate the transition function $\partial S'/\partial S$. The parameters of $\partial S'/\partial S$ are the strategies that govern the evolution of firms' portfolio characteristics; ultimately we are finding the cost parameters that make these strategies optimal by satisfying the first order condition (7). Since strategies are a function only of the current state vector, we can write:

$$S_{it+1} = g(S_{it}) + \epsilon_{it} \tag{8}$$

where $g(\cdot)$ is a transition function reflecting firms' strategies and ϵ_{it} is a vetor of iid shocks to strategy implementation, which can be thought of as reflecting adjustment cost shocks in a reduced form way.

Estimation of the transition equation (8) is difficult. The goal is to estimate the evolution of each element in $char_{it}$ and $char_{-it}$ as a function of all of the elements of both $char_{it}$ and $char_{-it}$.[26] It is crucial to identify the causal effect of this period's portfolio characteristic on next period's portfolio characteristic in order to produce unbiased estimates of the second-stage cost parameters. In general, though, regressing this period's state on last period's state introduces endogeneity problems. Most notably here, unobserved firm heterogeneity will bias estimates; for instance, it may appear that owning many stations this period causes a firm to own many stations next period, when really the most profitable firms tend to own a lot of stations in all periods. Also, time trends may bias estimates in the transition function. So, I would like to include firm fixed effects and a time trend in the specification of Equation (8), giving us:

$$\begin{gathered} char_{ijt+1} = \alpha_{0j}^{own} t + \alpha_{1j}^{own} char_{it} + \alpha_{2j}^{own} char_{-it} + \eta_{ij}^{own} + \epsilon_{ijt}^{own} \quad \forall \quad j = 1, \ldots, J \\ char_{-ijt+1} = \alpha_{0j}^{comp} t + \alpha_{1j}^{comp} char_{it} + \alpha_{2j}^{comp} char_{-it} + \eta_{ij}^{comp} + \epsilon_{ijt}^{comp} \quad \forall \quad j = 1, \ldots, J \end{gathered} \tag{9}$$

[26] The other elements of S_{it} are Δn_{it} which comes directly out of the evolution of $n_{it} \in char_{it}$, μ_i^r and μ_i^c which are constant, and ν_{it}, ω_{perit} and ω_{adjit} which are iid.

Notice that the α^{own}'s are the firm's own strategies, while the α^{comp}'s summarize its competitors strategies. By using this reduced-form linear specification for the transition of portfolio characteristics, I am making the assumption that firms can freely adjust their portfolio characteristics, and are not subject to integer constraints or restrictions imposed by the set of stations actually available for purchase.

Estimation of (9) still presents problems. In the presence of lagged dependent variables and endogenous regressors, η_{ij} cannot be treated as an estimable constant. First-differencing removes firm fixed effects, but the endogeneity problem requires instruments. I use the estimator developed by Arellano and Bond (1991), which uses higher-order lagged differences of the regressors as instruments. The following system of equations is estimated using the Arellano-Bond GMM estimator:

$$\Delta char_{ijt+1} = \alpha_{0j}^{own} + \alpha_{1j}^{own} \Delta char_{it} + \alpha_{2j}^{own} \Delta char_{-it} + \Delta\epsilon_{ijt}^{own} \quad \forall \quad j = 1, \ldots, J$$

$$\Delta char_{-ijt+1} = \alpha_{0j}^{comp} + \alpha_{1j}^{comp} \Delta char_{it} + \alpha_{2j}^{comp} \Delta char_{-it} + \Delta\epsilon_{ijt}^{comp} \quad \forall \quad j = 1, \ldots, J \tag{10}$$

The Arellano-Bond estimation method often suffers from a weak instrument problem. Fortunately the F-statistic is greater than 10 in all of the first-stage regressions but one, when it is just under 10. A bigger problem is that the validity of the instruments requires the assumption of no serial correlation, which is strong in this case. For each transition regression, I include enough lags such that the assumption of no serial correlation cannot be rejected. However, for computational reasons, I cannot use transition equations with numerous lags when estimating the first order conditions. Fortunately, the predictions do not change much when these additional lags are dropped.

Because I have specified linear transition functions, $\partial S'/\partial S$ and $\partial S'/\partial char_j$

are constant with respect to the state S. The first order conditions can therefore be written:

$$E\left[\sum_{\tau=0}^{\infty}\beta^{\tau}\frac{\partial\pi_{it+\tau}}{\partial S_{it+\tau}}'\left(\frac{\partial g}{\partial S}\right)^{\tau-1}\frac{\partial g}{\partial char_j}\mid S_{it}\right]=0\quad\forall\quad i,j,t \tag{11}$$

Estimation of the transition equation $g(cdot)$ gives us two components of the first order conditions (11); that is, $(\partial g/\partial S)^{\tau-1}$ and $\partial g/\partial char_j$. The next step is to find, for each state S_{it} observed in the data, the present discounted value of the stream of future marginal profits $E(\sum_{\tau=0}^{\infty}\beta^{\tau}\partial\pi_{it+\tau}/\partial S_{it+\tau}|S_{it})$ (which will then by multiplied by $(\partial g/\partial S)^{\tau-1}\partial g/\partial char_j$ period-by-period). Note from the revenue and cost equations (1) and (2) that the marginal effect of state variables on profits $(\partial\pi/\partial S)$ depends upon the number of stations owned; and in fact, $(\partial\pi/\partial S)$ may also depend on the values of other state variables if the state variables enter the per-station revenue or cost functions non-linearly.[27] Therefore we must evaluate $E(\partial\pi_{it+\tau}/\partial S_{it+\tau}|S_{it})$ at $E(S_{it+\tau}|S_{it})$ for each period $t+\tau$.

The expectation in $E(\partial\pi_{it+\tau}/\partial S_{it+\tau}|S_{it})$ is over shocks to demand (ν), cost (ω) and strategy implementation (ϵ). In order to estimate the expectation, I use the simulation method that Bajari et al. (2007) use to estimate value functions; here I am estimating the derivatives of the value functions, but otherwise the method is the same. For a given state S_{it}, a path of play can be simulated by using the estimated transition functions (10) and a set of shocks drawn from the estimated distributions of the ϵ's. I simulate the evolution of the state vector well into the future (100 periods), until the discount factor will render sufficiently small the present value of any marginal returns generated. Given a set of revenue and cost coefficients and draws of shocks ν and ω, I can calculate the present value of marginal returns associated with this path of play. Repeating

[27] I include quadratic terms in estimation.

this procedure numerous (1000-2000) times and averaging the present value of marginal returns over all of these paths gives me an estimated ex ante stream of marginal returns associated with this state S_{it}.

Because the industry remains regulated, firms are not free to choose any portfolio that they want. Therefore in simulations I restrict the strategies of firms accordingly: I do not allow firms to acquire more stations and achieve greater population coverage than would be allowed under current regulations, to choose more regional concentration than would be allowed, and so on. The limits on the portfolio characteristics are carefully chosen, and are allowed to depend on the firm's number of stations. There is no condition that forces a purchase of a station by one firm to be accompanied by a sale of a station by another firm. However, when I simulate paths of play using the estimated strategies, I can see whether future states are realistic; e.g., whether the total number of stations owned by all of the firms together stays roughly constant. This provides a nice check on the reduced-form estimation of strategies; in simulations, I find that the future states are indeed realistic.

6.3.2 Search for Cost Estimates

As explained above, for a given set of revenue and cost parameters, we can estimate the first order condition (11) by simulation. However, the simulation of numerous paths of play for each of the 1,799 states that are observed in the data is the computationally expensive part of the estimation process. Therefore I do not want to embed the simulations in the search for the cost coefficients, forcing me to re-do the simulations for every trial vector of cost coefficients. This can be avoided if marginal profits are linear in the cost coefficients, as explained below.[28]

[28]This is virtually equivalent to the linearity assumption proposed in Bajari et al. (2007).

Since static profit is simply revenue less cost, we have:

$$E\left[\sum_{\tau=0}^{\infty}\beta^{\tau}\frac{\partial R_{it+\tau}}{\partial S_{it+\tau}}'\left(\frac{\partial g}{\partial S}\right)^{\tau-1}\frac{\partial g}{\partial char_j}\mid S_{it}\right]$$

$$= E\left[\sum_{\tau=0}^{\infty}\beta^{\tau}\frac{\partial C_{it+\tau}}{\partial S_{it+\tau}}'\left(\frac{\partial g}{\partial S}\right)^{\tau-1}\frac{\partial g}{\partial char_j}\mid S_{it}\right] \quad \forall \quad i,j,t \tag{12}$$

That is, the present discounted value of the stream of expected marginal revenues with respect to $char_{ijt}$ must be equal to the present discounted value of the stream of expected marginal costs with respect to $char_{ijt}$. We have estimated the revenue equation and the transition equation, so we already have the lefthand side of Equation (12). If marginal costs are linear in the cost parameter vector c, then we can pull c out of the righthand side:

$$E\left[\sum_{\tau=0}^{\infty}\beta^{\tau}\frac{\partial R_{it+\tau}}{\partial S_{it+\tau}}'\left(\frac{\partial g}{\partial S}\right)^{\tau-1}\frac{\partial g}{\partial char_j}\mid S_{it}\right]$$

$$= c\cdot E\left[\sum_{\tau=0}^{\infty}\beta^{\tau}\psi(S_{it+\tau})'\left(\frac{\partial g}{\partial S}\right)^{\tau-1}\frac{\partial g}{\partial char_j}\mid S_{it}\right] \quad \forall \quad i,j,t \tag{13}$$

where $\psi(S_{it})$ is a basis function such that $\partial C_{it}/\partial S_{it} = c\cdot\psi(S_{it})$. We can run the simulations once, which simultaneously gets us the lefthand side and all of the righthand side except c; then we can estimate $\hat{c}$ by simply regressing the lefthand side on the righthand side. See the Appendix for a detailed discussion of how I estimate the lefthand and righthand sides by simulation.

The assumption that $\partial C_{it}/\partial S_{it} = c\cdot\psi(S_{it})$ is not as restrictive as it looks, because $\psi(S_{it})$ can include transformations of state variables, such as quadratic terms, interaction terms, etc. This means that I can easily allow for non-linearities in the relationships between portfolio characteristics and per-station profits. The main drawback is that I cannot use a log-linear specification for the cost equation.

I impose another restriction in order to make estimation feasible: the firm-specific cost constant μ_i^c is a linear function of the firm-specific revenue constant μ_i^r. In other words, I include μ_i^r as a variable in the cost equation and estimate the cost coefficient on it. This is much less computationally expensive than freely estimating μ_i^c, which would require the inclusion of 244 firm dummies (or transformations thereof) in ψ. By including μ_i^r as a regressor in the cost function, I am allowing for the possibility that firms that generate systematically higher revenue incur either systematically higher or systematically lower costs. This allows for two scenarios. In one scenario, some firms are simply "better," both generating higher revenue and incurring lower costs; this would lead to a negative relationship between $\hat{\mu}_i^r$ and μ_i^c. In another scenario, the firms that invest more in programming tend to draw in more advertising revenue, whereas firms that cut programming costs tend to take a hit to advertising revenue; this would lead to a positive relationship between $\hat{\mu}_i^r$ and μ_i^c. The cost coefficient on $\hat{\mu}_i^r$ tells us which scenario is a better description of reality. I do the same with the portfolio fixed effect: I include the revenue-side portfolio fixed effect (recall that this is the average revenue of the portfolio across all years, regardless of ownership) as a regressor in the cost equation, and estimate the cost coefficient on it.

7 Empirical Results

Revenue is observed, so the interpretation of revenue estimates is fairly straightforward. The interpretation of cost estimates is more complicated. Cost is essentially a residual, serving to explain the behavior of firms that is left unexplained by revenue. Therefore it is unreasonable to interpret cost simply as strictly the cash that flows out of a firm in a given year; it may include the cost of raising capital, the cost of being exposed to risk, and so on. This should be kept in mind when interpreting the results.

The first order conditions used to recover cost assume that the present discounted value of the expected stream of marginal revenue from a station equals the present discounted value of the expected stream of marginal cost of a station. Note that the annual marginal cost of owning an additional station includes the annual opportunity cost of holding (not selling) the station at market value.

Suppose that the market for broadcast television stations operates perfectly, with efficient credit markets and all agents having full information. Then we would expect the market value of a station to be approximately equal to the present discounted value of the station's expected stream of profits. Therefore the annual opportunity cost of holding the station would be approximately equal to the annual profit generated by the station. In this case, recovered annual per-station cost would be close to annual per-station revenue.[29] Table 2 presents summary statistics for annual per-station revenue and recovered annual per-station cost for portfolios that are observed in the data over the past ten years. The recovered cost numbers look reasonable. I find that 90% of firms in my sample[30] are earning a positive profit, suggesting that the market is not perfectly competitive.

It is worth thinking about how inclusion of opportunity cost in recovered cost may affect the estimated coefficients of interest, that is, the effects of a firm's portfolio characteristics on its per-station cost. As long as this opportunity cost

[29] Consider a simple numerical example. A station generates \$10 million in annual revenue, and costs \$7 million to operate annually; therefore the station's annual profit is \$3 million. If the market for this station operates perfectly, with full information, then the opportunity cost of holding the station would be approximately equal to \$3 million because this is how much other firms would be willing to pay annually in order to own the station. Therefore the *total* annual cost of the station is \$7 million + \$3 million = \$10 million.

[30] Recall that these are non-network firms owning more than one station.

is exogenous to the variables of interest, the estimated coefficients should not be biased. If the market for television stations is competitive, then the opportunity cost of holding a station should reflect the highest value that any *other* firm would pay for the station. This value should not be directly affected by the portfolio characteristics of the firm in question. Therefore the coefficients of interest should not be affected. However, if the market is not competitive, then the price of a station is determined through negotiation between the seller and the buyer. The price of the station will still be the highest value that any third party would pay for it (and the cost coefficients will be unbiased) only if buyers manage to extract all of the surplus from sellers during negotiations. If, instead, sellers are able to extract some surplus, then this will erode the cost advantage that the buyer has in the operation of that station; this would bias me *away* from finding statistically significant effects of portfolio characteristics on per-station cost.

A final point that should be kept in mind when interpreting my results is that this industry remains regulated. Therefore firms are not completely free to adjust the characteristics of their portfolios of stations. Therefore, the estimated revenue and cost effects should be interpreted as those that exist *under current regulations*. They do not tell us what the revenue and cost advantages of consolidation would be under an entirely different set of regulations.

The estimated effects of a firm's portfolio characteristics on its annual per-station revenue are shown in Table 3, and the effects on its annual per-station cost are shown in Table 4.[31] Note that the adjustment costs are one-time effects on the firm's total (not per-station) cost. The estimated strategic effects are shown in Table 5; these are the effects of the average portfolio characteristics

[31]Standard errors for the second-stage cost estimates have not yet been computed, but will be available soon.

among a firm's competitors on its own annual per-station revenue. In each table, I show the marginal effect as well as the effect of a standard deviation change in each portfolio characteristic; the latter is useful because the units of the portfolio characteristics vary and are not necessarily intuitive.

As shown in Table 3, I find that a standard deviation increase of 11.9 in *Total Population Coverage of Stations* (in millions of persons) causes a \$2.07 million increase in per-station revenue. That is, increasing the total population coverage of a firm's stations enables the firm to boost its per-station revenue, most likely because advertisers pay more to reach more viewers per transaction. Given that median per-station revenue is \$8.2 million, this effect is large in magnitude.

Per-station revenue is also significantly greater if a firm's stations are of similar sizes (in terms of population coverage), are located closer together (in miles) and offer diversity in terms of their network affiliations. A standard deviation decrease of 34.9 in the coefficient of variation in stations' coverage areas increases per-station revenue by \$1.2 million. A standard deviation decrease of 544 miles in the average distance between a firm's stations increase that firm's per-station revenue by \$1.2 million. And lastly, a standard deviation decrease of 1.0 in *Network Concentration* reduces per-station revenue by \$1.2 million. Most likely this is because an advertiser has the option of advertising through a national network rather than through the owner of network-affiliated stations. Therefore a firm's portfolio of stations is probably less attractive to an advertiser if it is essentially just a subset of the stations that are affiliated with one particular national network. This might explain why we see that firms tend to own a diverse set of stations in terms of their network affiliations. Also, it appears that there are no significant costs for a firm associated with an additional network affiliation: there is no effect of *Number of Networks* on per-station cost.

The revenue effects discussed so far reflect the ability of firms to charge higher advertising rates if they offer a portfolio that is attractive to advertisers. Aside from advertising, firms also earn revenue from fees from networks and cable providers. It is possible that larger firms have more bargaining power in negotations with networks and cable providers; this could be an additional advantage of consolidation in this industry. I find that bargaining power with network is not a major factor for firms; the effect of *Network Concentration* on per-station revenue is statistically insignificant. However, I find that if a firm owns a large fraction of cable providers' broadcast stations, this enables the firm to extract significantly more revenue: a standard deviation increase of 237.4 in *Cable Domination* raises a firm's per-station revenue by $1.2 million. This finding is supported by anecdotal evidence; there have been a number of stories in the news recently about broadcasting firms that are able to demand higher fees from cable providers, and even some speculation that this will become a major source of revenue in this industry, as broadcast stations struggle to maintain adequate advertising revenue.[32] Some of the disputes between broadcasting firms and cable providers have been high-profile. In March 2009, several cable providers asked the FCC to consider issuing rules regarding retransmission negotiations.

Thus the revenue-side advantages of consolidation appear to be significant in this industry. I also find evidence of significant cost savings associated with consolidation. The most basic measure of the potential for economies of scale is simply the number of stations a firm owns. I find that owning more stations is

[32]For instance, Nexstar Broadcasting publicly announced in January of 2006 that it would earn more than $40 million for its 41 stations from its retransmission deals with cable providers; by my measure, Nexstar's level of "cable domination" in 2006 was at the 95th percentile for the industry. Sinclair Broadcasting was expected to earn $20 million in retransmission fees for its 47 stations in 2006; by my measure, Sinclair's level of "cable domination" in 2006 was at the 99th percentile for the industry.

associated with significant cost savings: a standard deviation increase of 6.6 in *Number of Stations* owned is associated with a $414,000 reduction in per-station cost. Firms are evidently able to combine operations of multiple stations, even though these stations are in different markets, often located quite far apart; the average distance between a firm's stations is 741 miles. Interestingly, I do not find that firms enjoy greater economies of scale if their stations are closer together or concentrated regionally; however, it is important to keep in mind that firms are quite restricted in their ability to locate their stations close together, because it is difficult to own multiple stations per market and markets are, by construction, located far apart.

I find that the operation of multiple stations is less costly if the firms' markets are similar in terms of their sizes (coverage areas) and the demographics and incomes of their viewers. However, the magnitudes of these effects are not especially large. Standard deviation decreases in *Coverage: Coefficient of Variation*, *Demographic Heterogeneity* and *Income Heterogeneity* reduce per-station costs by $162,000, $111,000 and $ 82,000, respectively.

While I find that negotations with cable providers can yield large revenues, I find evidence that there are significant costs associated with these negotations. Specifically, I find that it is quite costly for a firm to deal with numerous cable providers: I find that a standard deviation increase of 18.9 in *Number of Cable Providers* raises per-station cost by $899,000.[33]

I find that there is some cost associated with fluctuations in a firm's revenue stream: a standard deviation increase of 15.2 in *Volatility of Revenue Stream*

[33] In each market in which a firm owns a station, there are a number of companies that provide cable service to viewers, and that therefore can retransmit the firm's station to viewers. (This number varies from one or two per market to many dozens.) The portfolio characteristic *Number of Cable Providers* is the number of *different* cable providers in all of the firm's markets.

increases per-station cost by $279,000. This is not likely to reflect actual cash flowing out of the firm; instead, this means that firms are willing to accept lower expected revenue in exchange for less volatile revenue. Another advantage of consolidation in this industry may be that owning a large number of stations (in different markets, affiliated with different networks, with different characateristics, and so on) helps firms smooth their revenue stream over time. As shown in Table 4, larger firms do achieve lower volatility in revenue by my measure.

The cost coefficient on the average revenue potential of a firm's stations is 1.06, meaning that stations that generate more revenue also cost more to operate (in fact, the extra cost slightly outweighs the extra revenue). The cost coefficient on the firm's revenue fixed-effect coefficient is 0.29. The positive sign means that firms that generate higher per-station revenue also incur higher per-station cost (though less so). It appears that firms that invest more in programming (increasing costs) tend to reap higher revenue in return, while firms that cut costs tend to accept revenue losses as a result. The intuition behind the identification of this parameter is as follows: firms that earn more revenue per station are not systematically seeking out large portfolios as much as we would expect given the higher per-station revenue that they generate; this must be because they also incur higher per-station costs.

While I find that there are significant advantages to consolidation in this industry, I also find that firms incur significant sunk costs when they acquire stations. I find that firms incur a sunk cost of $91,000 for every unit of squared positive increase in the number of stations owned; the effect of a standard deviation increase of 20.5 (a purchase of 4.5 stations) is $1.9 million in sunk costs. I actually find small negative sunk costs associated with selling a station, most likely reflecting the fact that selling a station can be an attractive way for a struggling firm to access cash.

I find that strategic effects are important in this industry; six of the thirteen strategic variables included in estimation have large and statistically significant effects on a firm's per-station revenue. A standard deviation increase of 6.1 in the average *Number of Stations* among competitors reduces per-station revenue by $414,000. This makes sense, given the finding that a firm faces lower per-station cost if it owns more stations. If a firm faces competitors with more stations, then the firm faces competitors with lower costs; its competitors can therefore offer lower advertising prices, and draws advertisers away. However, a standard deviation increase of 15.1 in the average *Total Population Coverage* among competitors actually increases per-station revenue by $3.3 million. This is a very large effect, and it is statistically significant at the 1% level. This may suggest that when firms cover large populations, they attract national or regional advertisers who are willing to pay a lot, and actually offer less competition for local advertisers.

8 Conclusion

This paper estimates the demand-side and supply-side advantages of consolidation in the broadcast television industry. The consolidation has provoked considerable controversy. However, the reasons it occurred are not well understood, because firms mainly purchase stations that are not in direct competition with one another and that are located many miles apart.

I exploit an exogenous change in regulation that led to significant consolidation in the industry. Although revenue and ownership data are available for broadcast television stations, cost data are not available. I recover costs by examining patterns in ownership changes that are left unexplained by revenue opportunities. Because the firm's purchasing decision involves dynamic and strategic elements, I model the decision as a dynamic game. The firm adjusts its portfolio characteristics each period so as to maximize the present discounted value of its expected stream of profits. My approach allows me to isolate the relationships of interest in an industry that each year witnesses hundreds of firms considering the purchase or sale of over one thousand stations. As far as I know, this paper is the first to estimate a model of merger activity in a dynamic, strategic framework.

My results suggest that consolidation in this industry offers significant revenue and cost advantages, working along a number of very different dimensions. First, I find that firms are able to raise more revenue per station if all of their stations together reach a larger total population, suggesting that firms can charge higher rates if they can offer advertisers more viewers per contract. I find that firms are better able to do this if their markets are closer together and similarly sized, and if they can offer stations with a variety of network affiliations. I also find evidence that a firm can extract higher fees from a cable provider if its stations represent a large percentage of the broadcast stations carried by the cable provider.

There are also cost-side advantages to consolidation. I find that firms face lower per-station costs if they own a greater number of stations, suggesting that firms are able to combine the operations of stations that are located in different markets. Firms can achieve greater economies of scale if their markets are similar in terms of their size, demographics and income levels. Consolidation also enables firms to reduce their exposure to revenue fluctuations.

The results suggest that a dynamic game is the appropriate framework to study this industry. I find that firms incur significant sunk costs when they purchase stations, suggesting that dynamics are important. Also, a firm's revenue is significantly affected by the portfolio characteristics of its competitors, suggesting that firms are likely to consider the strategic effects of their decisions.

The framework that I develop to study this industry could easily be applied to other settings. For instance, it would be natural to model the decisions of multi-product oligopolistic firms in this way; firms choose which products and/or product characteristics to include in their portfolios, taking into consideration the sunk costs involved in the introduction of new products and taking into consideration the strategic effects of their decisions. My framework allows for computationally inexpensive estimation of interesting relationships in this type of complex strategic environment in which firms make numerous interrelated decisions.

Appendix: Estimating expected first-order conditions by simulation

Coming soon!

References

Akkus, O. and Hortacsu, A. (2007). The determinants of bank mergers: A revealed preference analysis.

Arellano, M. and Bond, S. (1991). Some tests for specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58:277–297.

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

Berry, S. and Pakes, A. (2000). Estimation from the optimality conditions for dynamic controls.

Collard-Wexler, A. (2008). Demand fluctuations in the ready-mix concrete industry.

Crawford, G. (2007). Television station ownership structure and the quantity and quality of tv programming. Federal Communications Commission Media Ownership Study No. 3.

Fox, J. (2007). Estimating matching games with transfers.

Hammond, A., O'Connor, B., and Westin, T. (2007). The Impact of the FCC's TV Duopoly Rule Relaxation on Minority and Women Owned Broadcast Stations 1999-2006. Federal Communications Commission Media Ownership Study No. 8.

Ho, K. (2007). Insurer-provider networks in the medical care market. NBER Working Paper.

Hopenhayn, H. (1992). Entry, exit and firm dynamics in long run equilibrium. *Econometrica*, 60:1127–1150.

Jeziorski, P. (2010). Dynamic determinants of mergers and product characteristics in the radio industry.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, 50:649–670.

Kim, E. H. and Singal, V. (1993). Mergers and market power: Evidence from the airline industry. *The American Economic Review*, 83(3):549–569.

Klepper, S. (2002). Firm survival and the evolution of oligopoly. *The RAND Journal of Economics*, 33:37–61.

Klepper, S. and Grady, E. (1990). The evolution of new industries and the determinants of market structure. *The RAND Journal of Economics*, 21:27–44.

Klepper, S. and Simons, K. (2000). The making of an oligopoly: Firm survival and technological change in the evolution of the u.s. tire industry. *The Journal of Political Economy*, 108:728–760.

Knapp, W. (1990). Event analysis of air carrier mergers and acquisitions. *The Review of Economics and Statistics*, 72(4):703–707.

Park, M. (2008). An empirical two-sided matching model of acquisitions: Understanding merger incentives and outcomes in the mutual fund industry.

Paul, C. (2003). Cost economies: A driving force for consolidation and concentration? *Southern Economic Journal*, 70:110–127.

Pesendorfer, M. (2003). Horizontal mergers in the paper industry. *The RAND Journal of Economics*, 34(3):495–515.

Petrakis, E. and Roy, S. (1999). Cost-reducing investment, competition and industry dynamics. *International Economic Review*, 40:381–401.

Ryan, S. (2009). The costs of environmental regulation in a concentrated industry.

Shaked, A. and Sutton, J. (1983). Natural oligopolies. *Econometrica*, 51:1469–1483.

Shaked, A. and Sutton, J. (1987). Product differentiation and industrial structure. *The Journal of Industrial Economics*, 36:131–146.

Shaked, A. and Sutton, J. (1990). Multi-product firms and market structure. *The RAND Journal of Economics*, 21:45–62.

Shiman, D. (2007). The impact of ownership structure on television stations' news and public affairs programming. Federal Communications Commission Media Ownership Study No. 4: News Operations.

Snider, C. (2008). Predatory incentives and predation policy: The american airlines case.

Sorenson, M. (2007). How smart is smart money? a two-sided matching model of venture capital. *Journal of Finance*, 62:2725–2762.

Figure 1: Transactions Per Year

Figure 2: Distribution of Number of Stations Held Among Firms in 1995

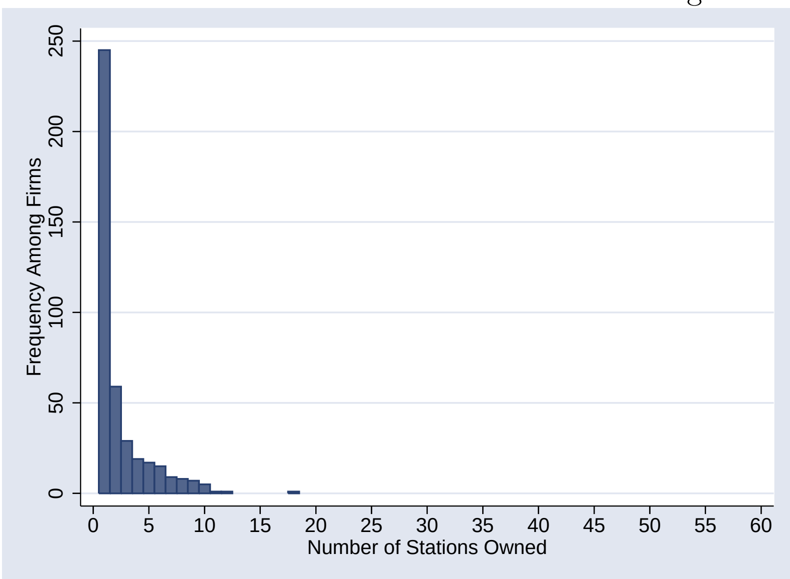

Figure 3: Distribution of Number of Stations Held Among Firms in 2007

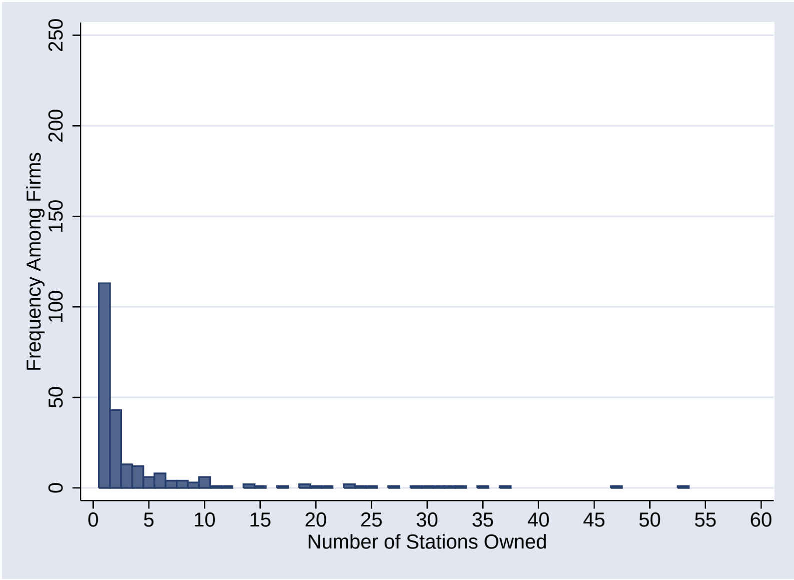

Figure 4: Expected Value of Coefficient of Variation in Revenue as Function of Number of Stations Owned

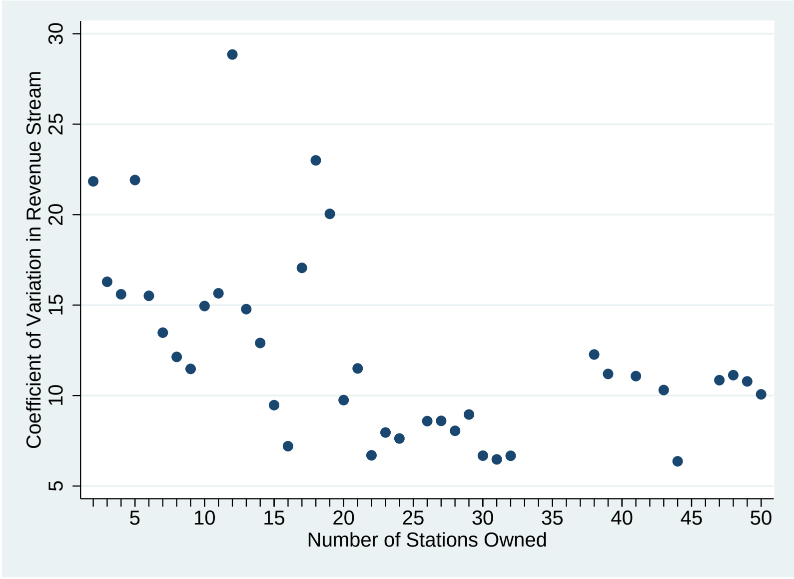

Table 1: Summary Statistics for Portfolio Characteristics

Variable	Median	Mean	Standard Deviation	Minimum	Maximum
Size					
Number of Stations	3	5.9	6.6	2	50
Total Populaton Coverage of Stations	2.5	6.8	11.9	0.02	102
Location					
Average Distance between Stations	630	741	544	0	3,580
Regional Concentration of Stations	1.24	1.4	0.55	0.51	4.14
Network Affiliations					
Network Concentration	1.65	1.74	1.00	0	5.82
Domination of Networks	0.61	3.00	6.44	0	68.42
Number of Networks	2	1.97	1.20	0	5
Cable Providers					
Domination of Cable Providers	69	155.6	237.4	0.02	1,683
Number of Cable Providers	14	19.8	18.9	1	127
Similarity of Markets					
Coverage: Coefficient of Variation	59	61.2	34.9	0.01	221.8
Demographic Heterogeneity	19.4	21.1	15.1	0	94.6
Income Heterogeneity	11	11.3	8.1	0	51.2
Risk					
Volatility of Revenue Stream	12.2	17.7	15.2	1.5	117.3

1,799 firm-year observations.

Table 2: Summary Statistics for Revenue and Cost

Descriptive Statistic	Annual Per-Station Revenue	Recovered Annual Per-Station Cost
Mean	$15.6 million	$11.1 million
Standard Deviation	$20.2 million	$18.4 million
Median	$8.2 million	$4.3 million
Minimum	$0.1 million	$0.07 million
Maximum	$141 million	$135 million

1,799 firm-year observations.
Revenue and Cost are in 2008 dollars.

Table 3: Effect of Own Portfolio Characteristics on Annual Per-Station Revenue†

Variable	Marginal Effect	Effect of SD Change
Size		
Total Populaton Coverage of Stations	173.87**	2,069**
	(71.15)	
Squared	0.35	50
	(0.63)	
Number of Stations	-43.97	-290
	(97.37)	
Squared	-2.87*	-125*
	(1.62)	
Location		
Average Distance between Stations	-2.23*	-1,213*
	(1.19)	
Squared	0.001	296
	(0.0003)	
Regional Concentration of Stations	458.91	252
	(1402.10)	
Squared	-337.30	-102
	(376.96)	
Network Affiliations		
Network Concentration	-1,314.38***	-1,314***
	(499.92)	
Squared	216.00**	216**
	(87.53)	
Domination of Networks	56.45	364
	(74.68)	
Squared	-0.92	-38
	(1.12)	
Cable Providers		
Domination of Cable Providers	5.04**	1,196**
	(2.31)	
Squared	-0.003*	-169*
	(0.0016)	
Similarity of Markets		
Coverage: Coefficient of Variation	-29.37*	-1,205*
	(16.67)	
Squared	0.18**	219**
	(0.08)	
Demographic Heterogeneity	2.59	39
	(43.57)	
Squared	0.37	84
	(0.67)	
Income Heterogeneity	19.26	156
	(81.04)	
Squared	-0.01	-1
	(2.43)	

†Revenue is in thousands of 2008 dollars.
* Statistically significant at 10% level.
** Statistically significant at 5% level.
*** Statistically significant at 1% level.

Table 4: Effect of Own Portfolio Characteristics on Annual Per-Station Cost†

Variable	Marginal Effect	Effect of SD change
Size		
Number of Stations	-174.08	-1,149
Adjustment Cost (for Positive Squared Δn)	91.31	1,869
Adjustment Cost (for Negative Squared Δn)	-13.44	-556
Location		
Average Distance between Stations	-0.10	-54
Regional Concentration of Stations	7.63	4
Network Affiliations		
Number of Networks	-7.47	-9
Adjustment Cost (Positive)	5.72	2
Cable Providers		
Number of Cable Providers	47.58	899
Adjustment Cost (Positive)	-7.49	-2
Similarity of Markets		
Coverage: Coefficient of Variation	4.63	162
Demographic Heterogeneity	7.38	111
Income Heterogeneity	10.17	82
Risk		
Volatility of Revenue Stream	18.35	279
Fixed Effects		
Station Fixed Effect Revenue Coefficient	1.06	20,213
Firm Fixed Effect Revenue Coefficient	0.29	1,232

†Except adjustment costs, which are one-time effects on total firm cost. Cost is in thousands of 2008 dollars. Second-stage standard errors have not yet been computed.
* Statistically significant at 10% level.
** Statistically significant at 5% level.
*** Statistically significant at 1% level.

Table 5: Effect of Competitor Characteristics on Annual Per-Station Revenue†

Variable	Marginal Effect	Effect of SD Change
Size		
Total Populaton Coverage of Stations	214.80***	3,252***
	(46.92)	
Number of Stations	-67.47**	-414**
	(34.04)	
Location		
Average Distance between Stations	0.49	102
	(1.05)	
Regional Concentration of Stations	274.86	67
	(371.58)	
Network Affiliations		
Network Concentration	97.61	40
	(118.00)	
Domination of Networks	-151.73***	-1,335***
	(36.10)	
Number of Networks	716.39***	394***
	(241.81)	
Cable Providers		
Domination of Cable Providers	4.55	723
	(3.87)	
Number of Cable Providers	-100.35	-1,746
	(332.24)	
Similarity of Markets		
Coverage: Coefficient of Variation	-14.51**	-211**
	(6.49)	
Demographic Heterogeneity	28.42	168
	(51.92)	
Income Heterogeneity	-110.18	-365
	(154.34)	
Risk		
Volatility of Revenue Stream	61.36***	279***
	(18.88)	

†Revenue is in thousands of 2008 dollars.
* Statistically significant at 10% level.
** Statistically significant at 5% level.
*** Statistically significant at 1% level.