Abstract: In this BULLETIN, using data on the prices paid by multichannel video programing distributors (“MVPDs”) for basic cable networks, I conduct a retrospective analysis of the price effects of the Comcast-NBCU merger. Estimates from both the difference-in-differences and lagged-dependent variable models indicate no systematic increase in the prices for Comcast’s networks following the merger, including general interest programming, news channels, and national and regional sports networks. Programming costs, however, exert a potent influence on affiliate prices, with full pass through in many cases. The evidence suggests either that there was no net positive effect on incentives to raise prices above competitive levels following the vertical merger, or else that the behavioral remedies placed on the Comcast-NBCU merger have been effective. Accordingly, excessive concern about the prices of programming following a vertical merger in the MVPD market appear unwarranted, at least when addressed by behavioral remedies.
I. Background

Unlike horizontal mergers, vertical mergers and acquisitions do not eliminate direct competitors and are widely held by economists and antitrust enforcers to be motivated principally by efficiency concerns. Further, vertical relationships have long been under the watchful eye of the Federal Communications Commission (“FCC”), the antitrust agencies, and Congress. For example, in the 1992 Cable Act, Congress codified the “program access rules,” which prohibit vertically-integrated multichannel video programming distributors (“MVPDs”) from discriminating against rivals for the price, terms and conditions of programming.

In 2009, the proposed merger of Comcast (the nation’s largest MVPD) and NBCUniversal (which owned and operated, among other things, both broadcast television stations and cable networks) refocused attention on the effects of vertical relationships in the multichannel video industry. The transaction was hotly contested by “consumer” advocacy groups and Comcast’s competitors, mainly on the grounds that Comcast would somehow seek to use its control of cable networks and broadcast stations as a weapon against its rivals. To address concerns about Comcast’s incentives to use its acquired assets to handicap competitors, the FCC imposed a

---


number of behavioral conditions on the merger. With respect to programming, Comcast was required to make available its programming to rivals at prices which were “the economic equivalent of the price, terms and conditions that a Similarly Situated MVPD would pay…” and in the event of a dispute, the FCC would apply “baseball-style arbitration” to establish a price. With such behavioral conditions (along with an assortment of unrelated “voluntary” merger commitments designed to assuage various political constituencies), the merger was eventually approved by both the FCC and the Department of Justice (“DOJ”) in 2011.

It is déjà vu all over again.

Concerns about anticompetitive actions by vertically-integrated MVPDs have arisen in the context of the proposed merger between AT&T and Time Warner, though resistance to this merger appears flavored with political intrigue. As proposed, the transaction would transfer both basic cable networks like CNN and subscription networks such as HBO from Time Warner to AT&T, stirring up fears that these networks could be used as weapons against AT&T’s MVPD

---

3 In the Matter of Applications of Comcast Corporation, General Electric Company and NBC Universal, Inc. For Consent to Assign Licenses and Transfer Control of Licensees, FCC 11-4, MEMORANDUM OPINION AND ORDER, 26 FCC Rcd 4238 (rel. January 20, 2011)(hereinafter “Comcast-NBCU Order”). It should be noted that the FCC has a long history of favoring behavioral remedies for vertical mergers. See, e.g., In the Matter of General Motors Corporation and Hughes Electronics Corporation, Transferees And The News Corporation Limited, Transferees, For Authority to Transfer Control, FCC 03-330, MEMORANDUM OPINION AND ORDER, 19 FCC Rcd 473 (rel. January 14, 2004) at ¶ 360 (“On the one hand, certain of the potential competitive harms inherent in vertically integrated programming/MVPD providers have been recognized as requiring special remedies to prevent potential abuses. On the other hand, the remedies chosen, at least in recent years, have not generally been structural remedies, such as prohibitions on common ownership of programming and distribution assets, but behavioral remedies, such as requirements for program access and nondiscrimination. This choice reflects the general recognition that vertical integration is less likely than horizontal integration to have anticompetitive effects and is more likely to promote efficiency.”)

4 Comcast-NBCU Order, id., Appendix A at Section 2.a.ii.

5 Comcast-NBCU Order, id. at ¶ 50.


and over-the-top competitors.\(^9\) As this merger involves no license transfers, the applicants did not have to seek FCC approval, so the FCC is in no position to impose behavioral remedies (or, extract extraneous conditions for unrelated political purposes).\(^10\) Nevertheless, Time Warner independently extended to third-party distributors the same sort of arbitration protections that the Government embraced in Comcast-NBCU merger.\(^11\)

The Department of Justice officially challenged the merger based in part on the fear that the merged entity might “demand higher prices and more favorable terms” from its rivals and seeks to enjoin the transaction.\(^12\) Somewhat surprisingly (in light of the Comcast-NBCU merger), in the case of the AT&T/Time Warner deal, behavioral conditions appear to be a non-starter for the DOJ. For example, media coverage of the merger prior to the filing of the DOJ’s Complaint indicates that the DOJ sought to impose structural remedies on the merger, including the divestiture of either “must have” programming assets (e.g., CNN) or DirectTV (AT&T’s satellite MVPD).\(^13\) Moreover, in a recent speech to the American Bar Association, Assistant Attorney General Makan Delrahim laid down a clear marker that he views behavioral conditions as “fundamentally regulatory” and “controversial” as they often “fail” to protect competition.\(^14\)

---


\(^14\) See, e.g., Assistant Attorney General Makan Delrahim Delivers Keynote Address at American Bar Association’s Antitrust Fall Forum, Washington, DC (November 16, 2017) (available at: https://www.justice.gov/opa/speech/assistant-attorney-general-makan-delrahim-delivers-keynote-address-american-bar) (“Our goal in remedying unlawful transactions should be to let the competitive process play out. Unfortunately, behavioral remedies often fail to do that. Instead of protecting the competition that might be lost in an unlawful merger, a behavioral remedy supplants competition with regulation; it replaces disaggregated decision making with central planning.”). Kwoka (2015) and Kwoka and Moss (2011) argue behavioral remedies have been
AT&T’s CEO Randall Stephenson, however, has made it publicly clear he has no intention of selling prized programming assets like CNN to close the deal.15

Vertical integration’s effect on programming prices is historically and today an interesting topic, and certainly one worthy of empirical analysis. In this BULLETIN, I conduct a retrospective analysis of the price effects of the Comcast-NBCU merger. Given the conditions on the merger directly related to programming pricing, my empirical analysis is perhaps best characterized as a test of the effectiveness of behavioral remedies. However, in light of assertions that behavioral conditions are ineffective, my results may also shed light on the claims that vertical relationships in the MVPD market lead to anti-competitive price increases. After all, if behavioral conditions are ineffective, then the pricing behavior must reflect the unrestrained incentives of Comcast.

Looking at the prices for basic cable programming—including general interest networks, news networks, and national and regional sports networks—I employ difference-in-differences and lagged dependent variable methods to test for price effects. Across all network types, I find no price increases (or decreases). Prices paid by MVPDs for Comcast’s networks are consistent with independently-owned networks. Whether the lack of post-merger price effects is the result of a lack of market power or behavioral remedies is unclear, but the evidence does suggest that simple behavioral remedies, if only as a back stop, may be sufficient to avail concerns regarding vertical mergers in the MVPD industry.

II. Background

In reviewing the Comcast-NBCU merger, the FCC was clear in its beliefs about the impacts of vertical integration of an MVPD into the programming market: “Comcast could use exclusionary program access strategies to reduce competition from all significant current and potential rivals participating in those markets.”16 In the MVPD market, Comcast competes with multiple rivals, including satellite providers, telephone companies offering MVPD services, new entrants like Google, and now over-the-top providers like Netflix.17 Almost all of these rivals


16 Comcast-NBCU Order, supra n. 3 at ¶ 39.

17 Id. at ¶ 43 (“Comcast competes with multiple MVPD rivals.”)
purchased, to a greater or lesser extent, Comcast’s and NBCU’s programming.\textsuperscript{18} As such, Comcast may raise the price of the content, or exclude access altogether, to handicap its rivals and, as a result, steal customers. The Commission concluded, “Comcast could obtain or (to the extent it may already possess it) maintain market power in video distribution, and charge higher prices to its video distribution subscribers than those consumers would have paid absent the transaction.”\textsuperscript{19} The Commission was particularly concerned about what it deemed to be “marquee programming,” such as national cable networks, regional sports networks, and local broadcast programming, asserting such programming lacked good substitutes and was essential to a MVPDs ability to compete.\textsuperscript{20}

\textbf{A. A Bargaining Model}

Fears that a vertically-integrated MVPD will use its control over cable networks to harm rivals follows from a simple Nash bargaining model discussed heavily during the Comcast-NBCU merger review.\textsuperscript{21} The Nash model predicts that the price, $P$, resulting from a bargain is determined by

\begin{equation}
P = \mu V + (1 - \mu)C,
\end{equation}

where $V$ is the value of the good to the buyer, $C$ is the production cost, and $\mu$ is the relative bargaining strength of the seller (where $0 \leq \mu \leq 1$). If $V$ is larger than $C$, then a bargain is possible, and the gains from a deal, $V - C$, are split according to the relative bargaining strength of the parties. In the simplest case, the bargaining strengths are equal ($\mu = \frac{1}{2}$) so the surplus is split evenly between buyers and seller.

A cost change, $\Delta C$, will be alter price according to the following expression,

\begin{equation}
\Delta P = (1 - \mu)\Delta C.
\end{equation}

The term $\Delta C$ need not reflect changes in the cost of production but may also reflect other changes in the profits of the firm. For instance, when a vertically-integrated firm sells inputs of production

\textsuperscript{18} Id. at ¶ 43.

\textsuperscript{19} Id. at ¶ 29.

\textsuperscript{20} Id. at ¶ 30.

to a rival, the resulting competition reduces the seller’s profit. This argument is the essence of the FCC and DOJ position on vertical integration in the MVPD industry. If Comcast sells its USA Network to a rival, then that rival takes customers from Comcast, reducing Comcast’s profits. Mathematically, the lost profit enters the bargain as a change in cost ($\Delta C$).

Naively, this opportunity cost has three components. First, if the rival does not have access to the network, it loses some share of its customers ($0 \leq d \leq 1$). Second, some portion of these customer losses switch to the vertically-integrated firm ($0 \leq \alpha \leq 1$). Third, the vertically-integrated firm earns $\pi$ profit per subscriber. Thus, the opportunity cost of selling to a rival is

$$\Delta C = \alpha d \pi = \Delta Q \pi,$$

(3)

Where $\Delta Q$ is the increase in the merged entity’s subscriber base if its rivals are forced to pay high prices for or are refused access to the merged entity’s programming. Generalizing a bit, we may rewrite Expression (2) as

$$\Delta P = \gamma (\Delta Q \pi).$$

(4)

where $\gamma$ is a bargaining parameter (equal to $1 - \mu$).22 During the Comcast-NBCU merger proceeding, there were numerous attempts to quantify these parameters, some more compelling than others. Still, it seems reasonable that the MVPD business is more complex than this simple model presumes, so that even good proxies for these parameters may render inaccurate predictions. Also, even if plausible values for $\alpha$, $d$, and $\pi$ may be estimated and the model a reasonable approximation of reality, the relative bargaining power of the two parties remains an open question, so that even if the opportunity cost is large the price effect may be small (if $\gamma$ is small).

How a merger effects prices is, in the end, an empirical question, and an important one. My objective in this BULLETIN is to apply accepted econometric techniques to data on affiliate prices in an attempt to quantify the actual price effects (rather than the predicted effects) of the Comcast-NBCU merger. As discussed next, however, it is difficult, if not impossible, to see clearly the incentives created by the merger in the data, since the FCC imposed numerous conditions on the merger. Today, where the effectiveness of behavioral remedies are in dispute, evidence on price effects under remedial constraint is also an important policy question.

---

22 Israel and Katz (2010), id.
B. Imposing Behavioral Remedies

Despite significant opposition, the Comcast-NBCU merger was approved, subject to extensive behavioral conditions. With respect to the concerns over programming, the FCC imposed an “arbitration remedy” that applied to all Comcast-NBCU programming. As it had done in prior transactions, the Commission imposed “baseball-style arbitration to maintain the pre-integration balance of bargaining power between vertically-integrated programming networks and rival MVPDs.” In baseball-style arbitration, the two parties submit prices and the arbiter chooses one or the other, thereby limiting the discretion of the arbiter. Under the rules, an aggrieved MVPD may initiate an arbitration at the Commission over the terms and conditions of carriage of any Comcast-NBCU affiliated programming. To further protect the MVPD, the MVPD is permitted to carry the programming that is the subject of arbitration until the dispute is resolved. As noted above, Time Warner has now extended to third-party distributors the same sort of arbitration protections that the Government embraced in Comcast-NBCU, thereby providing a direct nexus between the Comcast-NBCU and the AT&T-Time Warner mergers with respect to programming prices.

Naturally, it is sensible to ask why the FCC’s program access rules do not offer sufficient protection to rival MVPDs. According to the Commission, the program access rules address

---

23 Arbitration guidelines are provided in Comcast-NBCU Order, supra n. 3 at Appendix A.

24 Comcast-NBCU Order, id. at ¶ 50.


26 Comcast-NBCU Order, supra n. 3 at ¶ 50.

27 Supra n. 11.

28 It should be noted that the Commission has applied its program access-type conditions to vertically-integrated DBS providers in the past. See, e.g., In the Matter of General Motors Corporation and Hughes Electronics Corporation, supra n. 3 at ¶ 84 (the “program access rules (and other non-discrimination safeguards) serve several useful functions with respect to the video programming subject to the vertically integrated firm’s control. First, the program access rules prohibit permanent foreclosure with respect to all satellite cable programming. Second, they can prevent overt discrimination in the prices the integrated firm charges for such inputs. Finally, they can also prevent uniform increases in satellite cable programming input prices where the downstream affiliate is partially owned and where the cost of compensating the affiliate exceeds the expected profits resulting from the price increase.”); In the matter of News Corporation and the DirecTV Group, Inc., Transferors, and Liberty Media Corporation, Transferee, For Authority to Transfer Control, MEMORANDUM OPINION AND ORDER, FCC 08-66, 08 FCC Rcd 66 (rel. February 26, 2008) at ¶ 5 (“we require the Applicants to abide by program access, program carriage, RSN arbitration, and retransmission consent arbitration conditions modeled on the conditions imposed in the News Corp.-Hughes proceeding.”). Moreover, under the plain terms of Section 628(j), 47 U.S.C. § 548(j), “Any provision that applies to a cable operator under this section shall apply to a common carrier or its affiliate that provides video programming by any means directly to subscribers.” Thus, as
discriminatory pricing, thereby permitting a strategy of uniform price increases. A higher uniform price would not necessarily impact Comcast as the payments are an internal transfer. Therefore, the Commission concluded the Agency’s program access rules “inadequately address the potential harms presented by the increased ability and incentive of Comcast-NBCU to uniformly raise Comcast’s rivals’ fees.”\(^{29}\) Returning to the bargaining model, however, a uniform price increase is discriminatory, since the opportunity cost of the merged entity differs from that of the independent MVPD. The Commission ignored this fact.

C. The DOJ on the AT&T-Time Warner Merger

In its formal complaint challenging the AT&T-Time Warner merger, the DOJ co-opted the Commission’s arguments. The DOJ’s Complaint asserts, “the merged company would have the power to make its video distributor rivals less competitive by raising their costs, resulting in even higher monthly bills for American families.”\(^{30}\) While higher prices for Time Warner’s programming may reduce revenues for the merged entity, the Agency claimed that “[b]ecause the video distributor walking away from a deal with the merged company would lose access to [Time Warner]’s popular programming, some of the video distributor’s valuable customers would be dissatisfied and switch to a competing video distributor. Some of those departing customers would sign up with AT&T/DirecTV, bringing with them significant new profits for the merged company.”\(^{31}\) This “stealing business” effect “would give the merged firm the ability to credibly demand higher prices than it otherwise would.”\(^{32}\)

Naturally, the wording of the DOJ Complaint is forceful as it is preparing for litigation, but the arguments are no different than those offered by the FCC in the Comcast-NBCU merger. The DOJ’s argument flows directly from the simple Nash bargaining framework used in the Comcast-NBCU merger and outlined above.\(^{33}\) Studying the actual price effects of that merger is a worthwhile task.

---

\(^{29}\) Comcast-NBCU Order, supra n. 3 at ¶ 50.

\(^{30}\) DOJ Complaint, supra n. 12 at ¶ 3.

\(^{31}\) Id. at ¶ 5.

\(^{32}\) Id. at ¶ 5.

\(^{33}\) Comcast-NBCU Order, supra n. 3 at Appendix D.
III. Econometric Model

At the center of the FCC’s and DOJ’s argument regarding vertical mergers in the MVPD market is that the merged entity has an incentive to raise prices to its rivals, though there are tradeoffs to consider (e.g., the loss of advertising revenue). Thus, whether or not the merged company will raise programming prices above a competitive level is an open and empirical question. An empirical question requires an empirical answer. A retrospective analysis of the price effects of the Comcast-NBCU merger is a sensible place to start. The retrospective analysis of mergers has become popular in recent years, and serves as part of the empirical foundation of the “hipster” antitrust movement currently underway.

Given the conditions placed on the merger, it is impossible to evaluate the incentives of the merged firm from the constraints of the behavioral conditions. A decomposition of the incentives and the conditions is plausible under some outcomes. For instance, if prices rose sharply after the merger, then it may be presumed that the conditions were (mostly) ineffective, and the claims regarding the incentive to raise price valid. If prices fell, alternately, then either the predicted incentives are invalid or either the behavioral conditions held prices below those set under pre-merger incentives. If no effect is found, then either the merged firm is not incented (on net) to raise price or the behavioral conditions are effective. Thus, “no effect” is an informative result.

In a recent review of merger retrospectives, Kwoka (2015) described the differences-in-differences (“DiD”) regression as the “standard methodology for estimating the effects of mergers.” Numerous academic studies on merger effects—including (among many others) Tenn (2011), Ashenfeler and Hosken (2010), Ashenfelter, Hoskin, and Weinberg (2011), Taylor and Hosken (2007), Kwoka and Shumilkina (2010), and Thompson (2009)—apply the DiD

---

34 Id.
methodology to quantify the effects of mergers. As the standard approach, I turn to the DiD estimator to analyze the price effects of the Comcast-NBCU merger and its conditions.

The general format of the DiD regression is,

\[
\ln(P_i) = \delta D_{it} + \beta X_{it} + \lambda_i + \gamma_t + \epsilon_{it},
\]

where the dependent variable is the average price paid per subscriber month for the basic cable network \(i\) at time \(t\), \(D\) is a dummy variable equal to 1 for the treated group after the merger is completed, \(X\) are time-varying control variables (with coefficients \(\beta\)), \(\epsilon\) is the econometric disturbance term, and the \(\lambda\) and \(\gamma\) are the time and cross-section fixed effects (a two-way fixed effects model). The coefficient \(\delta\) is the average treatment effect (on the treated) that quantifies the price effects of the merger including any behavioral conditions that limit price increases.

Research indicates that the time series component of panel data (i.e., serial correlation) in DiD analysis systematically leads to the over-rejection of the null hypothesis. Unless otherwise indicated, statistical inference in this BULLETIN is based on clustered standard errors to address this concern. The efficiency of the clustering depends, in part, on the number of clusters, and in

---


a few instances the data used here contains “few” clusters (that is, few cable networks). Additional attention to statistical inference is provided in these cases.

The DiD estimator hinges on the validity of the parallel paths (or common paths) assumption, which holds that the difference between the treated and control groups are constant over time.\textsuperscript{40} If the assumption does not hold, then the DiD estimator is biased. The parallel paths assumptions cannot be formally tested, though the credibility of results are supported by the careful selection of the control group, perhaps using visual inspection of pre-treatment trends and/or conducting tests on pseudo-treatments. I apply both these methods.

O’Neill, Kreif, Grieve, Sutton and Sekhon (2016) propose an alternative estimation approach to DiD for calculating the treatment effect but one with good properties even when an appeal to the parallel paths assumption is weak.\textsuperscript{41} The Lagged Dependent Variable (“LDV”) model may be written as

\[
\ln(P_t) = \delta D_i + \beta X_{it} + \sum_{k=1}^{n} \theta_k \ln(P_{it-k}) + \nu_{it} \quad \forall t > T_0, \tag{6}
\]

where, if Equation (6) represents the true data generation process, \( D_i = 0 \) is the counterfactual outcome and, thus, \( \delta \) is the average treatment effect on the treated (“ATT”). Equation (6) may be estimated by Ordinary Least Squares (“OLS”) with only post-treatment observations \((T_0)\). A downside of this approach is the often substantial reduction in sample size.

Conceptually, the lagged dependent variables are intended to serve as proxies for unobserved influences, which are captured by the fixed effects in Equation (5). Thus, the LDV approach is best suited for problems with many pre-treatment observations, and O’Neill, et al., (2016) recommend exhausting all pre-treatment observations. Monte Carlo analysis in O’Neill, et al., (2016) demonstrates that across multiple estimation methods, including DiD, the LDV approach


produces the most efficient and least biased estimates when the parallel paths assumption is violated. The simulations confirmed that DiD performs quite well when the parallel paths assumption holds.

IV. Data

S&P Global’s Market Intelligence database, which is the best in class for cable network economics, is the source for all data used here. The dependent variable is the average affiliate revenue per subscriber month, which is the average price paid by MVPDs per subscriber-month for a basic cable network. For example, in 2016 the MVPDs paid about $0.09 per subscriber-month for the Cooking Channel, about $0.70 for Nickelodeon, and $7 for ESPN.

The Comcast-NBCU merger was consummated in 2011, so I define the treatment period as years 2012 through 2016 (five years) or 2013 through 2016 (four years). The data is annual. As is standard, I exclude 2011 or both 2011 and 2012 as a transition period. The pre-treatment period is selected as 2001 through 2010 (ten years). Some sample limitations are imposed at the beginning. First, networks that were shuttered prior to 2016 are dropped from the sample. Second, a comparison between the pre- and post-treatment periods requires sufficient data in both periods, so any networks lacking at least five years of data during the pre-treatment period are excluded from the sample.

42 The database is a subscription service (https://www.snl.com).
43 Data on prices paid by individual MVPDs per channel is not available.
44 Kwoka and Shumilkina (2010), supra n. 37.
45 Comcast’s G4 network was shutdown in 2014. D. Cowan, G4 Channel Shutting Down for Good this Month, ENGADGET (November 17, 2014) (available at: https://www.engadget.com/2014/11/17/g4-channel-shutting-down-for-good-this-month).
46 These data-based limitations on the sample affect only a few networks, and none of the major networks, so are of little consequence.
Table 1. Comcast-Owned Basic Cable Networks
(Millions, 2016)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Interest</strong></td>
<td></td>
<td></td>
<td><strong>National Sports</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bravo</td>
<td>89</td>
<td>226</td>
<td>Golf Channel</td>
<td>77</td>
<td>84</td>
</tr>
<tr>
<td>Cloo</td>
<td>24</td>
<td>29</td>
<td>NBA TV</td>
<td>53</td>
<td>43</td>
</tr>
<tr>
<td>E!</td>
<td>91</td>
<td>207</td>
<td>NBCSN</td>
<td>83</td>
<td>121</td>
</tr>
<tr>
<td>Esquire</td>
<td>63</td>
<td>...</td>
<td>NFL Network</td>
<td>71</td>
<td>117</td>
</tr>
<tr>
<td>Oxygen</td>
<td>77</td>
<td>109</td>
<td>Olympic Channel</td>
<td>30</td>
<td>...</td>
</tr>
<tr>
<td>Syfy</td>
<td>91</td>
<td>267</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>93</td>
<td>646</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>News</strong></td>
<td></td>
<td></td>
<td><strong>Regional Sports</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CNBC</td>
<td>90</td>
<td>144</td>
<td>NBC-Bay Area</td>
<td>4.0</td>
<td>22.5</td>
</tr>
<tr>
<td>CNBC World</td>
<td>36</td>
<td>...</td>
<td>NBC-Boston</td>
<td>3.7</td>
<td>6.6</td>
</tr>
<tr>
<td>MSNBC</td>
<td>81</td>
<td>481</td>
<td>NBC-California</td>
<td>3.4</td>
<td>3.6</td>
</tr>
<tr>
<td>Weather Channel</td>
<td>87</td>
<td>155</td>
<td>NBC-Chicago</td>
<td>4.6</td>
<td>15.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NBC-Philadelphia</td>
<td>2.8</td>
<td>13.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NBC-Washington</td>
<td>4.2</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Source: S&P Global, Market Intelligence.

Basic cable networks owned by Comcast after the merger cross many genres including broadcast television, general variety programming (e.g., Bravo and USA), national sports programming (e.g., NBCSN and NFL Network), regional sports networks (e.g., NBC Sports Boston) and news programming (e.g., CNBC and MSNBC). The networks are listed in Table 1. When estimating the price effects of the merger, I treat these different genres separately.47

Due to data limitations, broadcast networks unfortunately are excluded from the analysis. Premium channels (e.g., HBO, Showtime) are also excluded from the analysis as the merger did not involve such networks. Some networks acquired by Comcast in the merger were subsequently sold (e.g., A&E), and these properties are also excluded from the sample.48 In all, the data includes 22 networks owned by Comcast after the merger, with seven general interest networks, four news networks, five national sports networks, and six regional sports networks.49

47 Included in the general variety-programming category are Kagan’s genre classifications Arts & Entertainment, General/Variety, Women’s, Niche Networks, and Film.


V. Results

Difference-in-differences estimation compares the price changes of comparable firms, products or services before and after a treatment. Table 1 indicates significant differences among networks even within a single genre. In some cases, my analysis further disaggregates the genres and, for reasons explained below, excludes some networks from the analysis. Some the exclusions are based on the desire to satisfy the parallel paths assumption that is essential to the validity of the DiD estimator (and the efficiency and power of the LDV model). While the parallel paths assumption is formally untestable, there are some methods that lend credibility to the control group’s role as proxy for the untreated outcomes of the treated group. I employ three approaches. First, I employ visual inspection to detect clear discrepancies and anomalies in the pre-treatment pricing trends of the networks (see Figure 1 below). Second, I test for differences in the growth rates of the dependent variable between the treated and control group during the pre-treatment period.\textsuperscript{50} Third, I create a pseudo treatment for years 2008 through 2010, the last three years before the merger. Ideally, the treatment effect should not be statistically different from zero if the parallel paths assumption is valid.\textsuperscript{51} Networks dropped from the sample from visual inspection and the results of these two tests are detailed below.

As prices reflect costs, I include as a regressor in all models (the natural log of) programming expenses (with coefficient $\beta$ in Eqs. 5 and 6). As for other factors, the DiD regression includes cross-section fixed effects, thereby accounting for unobserved variables that do not vary over time. The time fixed effects account for changes in conditions common to all networks over time. All prices are converted to real dollars using the GDP Deflator with 2016 chosen as the base year.\textsuperscript{52} Any price changes are measured in real dollars. For the LDV model, fixed effects are not used; instead, the lagged dependent variables are used to quantify unobservable factors.\textsuperscript{53} General Interest Programming

For comparison purposes, I begin with the estimation of the price effects using all the general interest networks. There are 55 networks in the sample, seven of which are owned by Comcast. Results are provided in Table 2. The DiD estimates are based on either the exclusion of 2011 or 2011 and 2012 from the sample as transitions years. As for the parallel paths assumption, I am

\textsuperscript{50} This test involves regressing the dependent variable on a time trend and testing for differences across the control and treatment groups.

\textsuperscript{51} Pseudo-treatment may also be used to evaluate unconfoundedness. See Imbens and Wooldridge, supra n. 36.

\textsuperscript{52} ECONOMIC RESEARCH: FEDERAL RESERVE BANK OF ST. LOUIS, GDP Implicit Price Deflator in United States (available at: https://fred.stlouisfed.org/series/USAGDPDEFAYM1).\textsuperscript{53}

\textsuperscript{53} O’Neill, et al. (2016), supra n. 41.
unable to reject the null hypothesis of equal growth rates between the treated and control groups during the pre-treatment period or the null hypothesis of a zero pseudo-treatment effect. Visual inspection, however, indicates a number of networks are unlikely to serve as valid controls.

Table 2. General Interest, All Networks

<table>
<thead>
<tr>
<th>Model</th>
<th>( \delta )</th>
<th>( \beta )</th>
<th>N</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiD (2011 excl.)</td>
<td>0.039</td>
<td>0.511***</td>
<td>794</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>(0.35)</td>
<td>(5.85)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DiD (2011-12 excl.)</td>
<td>0.045</td>
<td>0.515***</td>
<td>739</td>
<td>0.493</td>
</tr>
<tr>
<td></td>
<td>(0.38)</td>
<td>(5.91)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDV</td>
<td>0.017</td>
<td>0.525***</td>
<td>241</td>
<td>0.855</td>
</tr>
<tr>
<td></td>
<td>(0.31)</td>
<td>(6.59)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Network clustered t-statistics in parentheses. Stat. Sig. * 10%, ** 5%, *** 1%.

Recognizing the limitations of this approach, Table 2 reveals that the null hypothesis of a zero price increase from the merger or its conditions cannot be rejected at anywhere near standard significance levels for either the DiD or LDV models. The sample sizes and the R² values are large. While the merger has no effect on prices, programming costs do. A 10% increase in programming costs translates into a roughly 5% increase in prices. As programming costs are 48% of revenues for this sample, the estimated coefficient on costs indicates approximately a 100% pass through of cost changes to price. Economic theory demonstrates that the pass through rate is not an indicator of competitiveness, but the pass through rate is likely of some policy significance. Also note that the Arellano-Wooldridge test for random effects rejects that specification in favor of fixed effects, and the (robust) F-Test for “no fixed effects” is rejected at the 1% level.

The estimates of the price effects of the merger may be improved by narrowing the analysis to more similarly-situated networks. First, I exclude the Cloo network given its relatively small audience size and subscription base. Second, while Oxygen has a larger audience and subscription base than Cloo, it is relatively small compared to the rest of Comcast’s general interest programming, being about half that of the next largest Comcast network (E!). Oxygen is

---


55 The Sargan-Hansen statistic (distributed \( \chi^2 \)) is 66.1 with a probability < 0.01. M. Arellano, *On the Testing of Correlated Effects with Panel Data*, 59 JOURNAL OF ECONOMETRICS 87-97 (1993); J.M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data* (2002) at pp. 290-1. A robust F-test (for generalized heteroskedasticity) is used to test the fixed effects (\( F = 1.4e+05, \) prob < 0.01). On the robust F-test, see also Cameron and Miller, supra n. 39. Similar results obtain for the restricted samples discussed later in the text.
excluded from the analysis. Third, among the Comcast networks, the USA network has by far the largest audience and is one of the nation’s most popular cable networks. Thus, I group the general interest networks into two groups based on network viewership: moderately sized (Bravo, E!, Syfy) and large sized (USA).

1. **Moderately Sized, General Interest Networks**

Comcast’s Bravo, E!, and Syfy networks are moderately sized with average audiences of just over 200 million viewers. A control group for these three networks is chosen from general interest networks with similarly sized audiences (160 million to 320 million in 2016, roughly a 50 million viewership boundary around the treated group). This selection criterion produces thirteen possible control networks. With respect to the parallel paths assumption, visual inspection of pre-treatment trends leads to the elimination of four networks, for a final control group of nine networks. In the pre-treatment period, growth rates are equal and the pseudo-treatment is statistically zero.

Without belaboring the point, the purpose of visual inspection is illustrated in Figure 1. The figure shows the (centered) pre-treatment pricing trends for the three Comcast networks and “OWN: The Oprah Winfrey Network.” As shown clearly in the figure, the pre-treatment trend for OWN would not satisfy the parallel paths assumption. Including it as a control would almost certainly lead to poor estimates of the treatment effect. OWN, along with three other networks

---

56 The control group includes: Animal Planet; Comedy Central, National Geographic, Spike, TLC, Travel Channel, WE tv, WGN America, and truTV. Dropped from the sample due to visual inspection are: GSN, LMN, OWN, and Science, all of which exhibit highly variable price changes uncharacteristic of the other networks in the group during the pre-treatment period.
that exhibit price trends unlike that of the Comcast networks, are excluded from the control group. There are twelve networks in the final group.

<table>
<thead>
<tr>
<th>Table 3. General Interest, Moderate Sized Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>DiD (2011 excl.)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>DiD (2011-12 excl.)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>LDV</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Network clustered t-statistics in parentheses. Stat. Sig. * 10%, ** 5%, *** 1%.

Results from the DiD and LDV models for the moderately-sized, general interest networks are summarized in Table 3. As before, across the three models none of the price changes are found to be statistically different from zero. Point estimates are small and vary in sign. Costs are again statistically powerful determinants of prices, though only for the LDV model and the coefficients are much smaller for this sample. Programming costs are about 36% of revenues for this group.

There are twelve networks in this particular sample, which is a number close to “few” in the context of clustered standard errors. Normally, having too few groups exacerbates the over-rejection problem, but the null hypothesis on the merger dummy is not rejected for any of the models in Table 3. The same is true if the clustering is set aside—the treatment variable is not statistically different from zero. As expected, clustering increases the standard errors and, consequently, reduces the t-statistics.

2. Large, General Interest Networks

Comcast’s USA Network is one of the most popular channels in the nation, with an audience size nearly triple that of Comcast’s Bravo network. A control group for the USA network is chosen from general interest networks with audiences of greater than 400 million viewers. Of the six possible control networks satisfying this criterion (including obvious controls like TNT and TBS), visual inspection and growth rate analysis leads to the elimination of one network. Using

---


58 The control group includes: Discovery, Food Network, HGTV, TBS, and TNT. Dropped from the sample due to visual inspection is Investigation Discovery.
these five controls, the pre-treatment growth rates are equal and the pseudo-treatment is statistically zero.

<table>
<thead>
<tr>
<th>Table 4. General Interest, USA Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>DiD (2011 excl.)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>DiD (2011-12 excl.)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>LDV</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Network clustered t-statistics in parentheses. Stat. Sig. * 10%, ** 5%, *** 1%.

Results for the larger networks are summarized in Table 4. Across the three models, the point estimates are all negative, but none are statistically different from zero (though most are in a one-tailed t-test). Costs are statistically significant in all models, but the coefficients are very different. Programming costs are about 30% of revenue for these networks, so the pass through rate is quite high on average, exceeding 100% in the DiD models but being nearly equaling 100% in the LDV model. Note that the sample size for the LDV model is only 30 observations, a result of the exclusion of pre-treatment observations.

There are only six networks in this sample; a small number of groups for clustering. The small number of networks also produces large variations in the coefficients across model types. Still, the \( \delta \) coefficient is statistically insignificant in all cases, so the problem of over-rejection of the null hypothesis does not appear to be an issue. If clustered errors are not used, then the only material change is that the \( \beta \) coefficients are all statistically significant at the 1% level.

The USA Network’s pre-treatment trends are very similar to those for the other Comcast channels. In light of the small number of networks in the USA Network analysis, in Table 5 I summarize the results when pooling the moderately-sized and USA Network samples used above (19 networks in all). The results are comparable, and the pre-treatment trends are equal and the pseudo-treatment is statistically zero. The null hypothesis that the \( \delta \) coefficient is zero is never rejected at standard significance levels and is negative in all cases.\(^{59}\) The \( \delta \) and \( \beta \) coefficients are also similar across model types.

\(^{59}\) The larger network count allowed consideration of two other networks with pre-treatment trends most unlike the control group (Food Network and HGTV). The results were materially the same.
Table 5. General Interest, Moderate Sized & USA Network

<table>
<thead>
<tr>
<th>Model</th>
<th>δ</th>
<th>β</th>
<th>N</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiD (2011 excl.)</td>
<td>-0.033</td>
<td>0.211**</td>
<td>285</td>
<td>0.739</td>
</tr>
<tr>
<td></td>
<td>(-0.32)</td>
<td>(1.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DiD (2011-12 excl.)</td>
<td>-0.032</td>
<td>0.207**</td>
<td>266</td>
<td>0.746</td>
</tr>
<tr>
<td></td>
<td>(-0.31)</td>
<td>(1.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDV</td>
<td>-0.046</td>
<td>0.292***</td>
<td>95</td>
<td>0.969</td>
</tr>
<tr>
<td></td>
<td>(-1.29)</td>
<td>(4.10)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Network clustered t-statistics in parentheses. Stat. Sig. * 10%, ** 5%, *** 1%.

The results for the general interest networks follow a clear pattern—the Comcast-NBCU merger has had no special effect on affiliate prices for its networks. Across different models and samples, the price effects are not statistically different from zero, and the coefficients are often small and in many cases negative. These results suggest that either the vertical merger between distribution and content has had no net positive pricing incentives, or that the behavioral conditions placed on that merger have held price increases in check. Affiliate prices, however, are found to be heavily influenced by programming costs.

D. News Networks

Turning now to news programming, the data includes ten networks, four of which are owned by Comcast. Estimating a meaningful DiD estimate of price effects for the news networks faced numerous challenges. Many of these networks—such as C-SPAN, Newsy, The Weather Channel—are quite specialized, and their price trends are highly variable. Also, affiliate fees for Bloomberg, Fox News and CNBC World (a Comcast property with a relatively small subscriber base and audience) rose sharply during the pre-treatment period, a pattern unlike the other networks in this genre. Unfortunately, given the characteristics of the news networks, treatment and control groups could not be acceptably constructed. As such, I believe the treatment effects are certain to be poorly estimated.

Table 6. News Networks

<table>
<thead>
<tr>
<th>Model</th>
<th>δ</th>
<th>β</th>
<th>N</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiD (2011 excl.)</td>
<td>-0.068</td>
<td>0.406</td>
<td>145</td>
<td>0.653</td>
</tr>
<tr>
<td></td>
<td>(-0.26)</td>
<td>(1.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DiD (2011-12 excl.)</td>
<td>-0.046</td>
<td>0.398</td>
<td>135</td>
<td>0.649</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(1.78)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDV</td>
<td>0.066</td>
<td>0.368</td>
<td>45</td>
<td>0.945</td>
</tr>
<tr>
<td></td>
<td>(0.48)</td>
<td>(1.47)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Network clustered t-statistics in parentheses. Stat. Sig. * 10%, ** 5%, *** 1%.
With that caveat in mind, Table 6 summarizes the results when all the news channels are included in the estimation sample. Surprisingly, neither of the checks on the parallel paths assumption indicate a problem—the growth rates are equal and the pseudo-treatment is zero. These results may simply reflect the high variability of the pre-trend prices across the networks in the sample. Visual inspection reveals significant disparity across the networks. As shown in the table, the null hypothesis of “no effect” of the merger cannot be rejected in any of the models, but I assign little credibility to the results for the news category. Costs are not statistically significant in the two-tailed test, but marginally significant in the one-tail test, especially for the DiD models. The revenue-programming cost margins for the news networks are very low (programming costs are 85% of revenues), so the β coefficients suggests about a 50% pass through rate.

E. National Sports Networks

From an empirical perspective, national sports networks are not unlike news networks in many regards. There are relatively few—nineteen in all—with some having broad coverage (e.g., ESPN) and others focusing on a single sport (e.g., The Golf Channel). Visual inspection of the pre-treatment trends indicates significant variation among the networks, complicating the choice of the control and the treated group. NBA TV, the Olympic Channel (both Comcast properties), Fox Sports Desportes, Fox Sports 1 and the Outdoor Channel and the Sportsman Channel all had nearly flat or even declining prices during the pre-treatment period (a rarity in this group). On the other hand, ESPN News, GolTV, and the NFL Network (the latter being a Comcast property) had rapidly rising prices during the pre-treatment period. Matching the networks into a relatively homogeneous group was not possible, so I exclude these oddities. In the final sample, there are four treated and eight control networks.

<table>
<thead>
<tr>
<th>Model</th>
<th>δ</th>
<th>β</th>
<th>N</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>DiD (2011 excl.)</td>
<td>-0.023</td>
<td>0.205*</td>
<td>138</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>(-0.29)</td>
<td>(2.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DiD (2011-12 excl.)</td>
<td>-0.022</td>
<td>0.206*</td>
<td>128</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>(-0.27)</td>
<td>(2.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LDV</td>
<td>-0.043</td>
<td>0.240*</td>
<td>38</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(-0.62)</td>
<td>(1.89)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Network clustered t-statistics in parentheses. Stat. Sig. * 10%, ** 5%, *** 1%.

Results for the national sports networks are summarized in Table 7. All the estimated δ coefficients are negative, though the treatment effects are not statistically different from zero. The analysis of the national sports networks provide additional evidence that the inventive to raise prices appears absent, or else the behavioral conditions placed upon the merged firm are
effective. Costs are a significant determinant of price, with a price increase of about 2% for every 10% cost increase. On average, programming costs are about 58% of revenue for these networks.

F. Regional Sports Networks

Comcast owns and operates six regional sports networks under the “NBC Sports” brand. A sensible control group is the thirteen regional sports networks owned and operated by Fox. Visual inspection recommends dropping a few of the regional networks. Also, including all of these networks leads to statistically significant differences in the pre-treatment growth rates and pseudo-treatment effect. Excluded from the estimation sample are Fox Sports Arizona, Detroit, and Southeast channels, and NBC Sports networks for Boston and Washington. Thirteen networks are in the final sample, with nine networks in the control group. For the filtered sample, the growth rates are statistically equal and pseudo-treatment is statistically indistinguishable from zero.

<table>
<thead>
<tr>
<th>Table 8. Regional Sports Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>DiD (2011 excl.)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>DiD (2011-12 excl.)</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>LDV</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Network clustered t-statistics in parentheses. Stat. Sig. * 10%, ** 5%, *** 1%.

The results for the regional sports networks are summarized in Table 8. The point estimates from the DiD models are quite small. None of the treatment effects is statistically different from zero. Prices are influenced by costs, with about half of cost increases passed through to price. As programming costs are about 57% of revenues, the results indicate that the regional sports networks fully pass on cost changes to affiliate prices.

Regional sports networks are frequently claimed to be an area where a vertically-integrated firm may exercise pricing power. While commenters expressed a fear that “Comcast’s dominance of regional sports programming” may result in higher prices, the Comcast-NBCU merger appears

---

60 The pre-treatment growth rates are identical for both samples, and the pseudo-treatments are likewise zero. Including some, but not all, of the excluded networks often leads to the rejection of the null hypothesis on the pre-treatment tests.
to have no effect on the prices for Comcast’s regional sports networks, either from a lack of incentive or the constraint of the behavioral conditions.\textsuperscript{61}

VI. Conclusion

Vertical relationships in the multichannel video programming delivery business have long been under the watchful eye of the Federal Communications Commission (“FCC”), the antitrust agencies, and Congress. Despite this oversight (and, more importantly, scant evidence of vertically-integrated firms abusing their control over programming prices to harm rivals\textsuperscript{62}), some continue to argue that vertically-integrated firms have the ability and incentive to abuse their control over programming prices to harm rivals and, therefore, draconian structural remedies are required. To test this argument, I conduct a retrospective analysis of the price effects of the Comcast-NBCU merger using data on the prices paid by multichannel video programing distributors (“MVPDs”) for basic cable networks. Estimates from both the difference-in-differences and lagged-dependent variable models indicate no systematic increase in prices for Comcast’s networks following the merger, including general interest programming, news channels, and national and regional sports networks. Programming costs, however, exert a potent influence on affiliate prices, with full pass through in many cases. The evidence suggests either that there was no net positive effect on incentives to raise prices above competitive levels following the vertical merger, or that the behavioral remedies placed on the Comcast-NBCU merger have been effective. Accordingly, excessive concern about the prices of programming following a vertical merger in the MVPD market appear unwarranted, at least when addressed by behavioral remedies.

\textsuperscript{61} Comcast-NBCU Order, supra n. 3 at ¶ 133.

\textsuperscript{62} The only direct evidence on the effects of programming prices is G.S. Ford and J.D. Jackson, Horizontal Concentration and Vertical Integration in the Cable Television Industry, 12 REVIEW OF INDUSTRIAL ORGANIZATION 501-518 (1997), though it does not assess the price effects of a merger or discriminatory pricing for individual channels.