

# **Market Power & Reciprocity Among Vertically Integrated Cable Providers**

**Jeffrey (Shih-kai) Shen**

*Professor Leslie M. Marx, Faculty Advisor*

---

*Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in  
Economics in Trinity College of Duke University.*

Duke University  
Durham, North Carolina  
2011

## **Acknowledgements**

I would like to thank Dr. Leslie Marx for her guidance, without which this project would not have been possible. Also I would like thank Dr. Michelle Connolly for her support and advice. I would like to thank Dr. Greg Crawford for his time and assistance. Lastly I would like to thank Tamara Louie for her edits and supporting me through the whole process.

---

JEL classification: C01, D22, K21.

Keywords: Vertical Integration, Cable Provider, Programming Network, Programming Distributor, Empirical Analysis.

## **Abstract**

This paper seeks to investigate the effects of vertical integration on the cable industry. There are two main goals that the research paper will attempt to address. The first is to build upon existing research on favoritism shown by multichannel video programming distributors (MVPDs) to affiliated video programming networks. Second, the paper will use 2007 and 2010 industry data to investigate the possible existence of “quid pro quo” among vertically integrated MVPD cable providers. After evaluating the data with multivariate OLS Regressions, the evidence suggests that MVPD cable providers do tend to carry their own affiliated programming networks. Furthermore, the evidence supports the hypothesis that reciprocity relationships exist among major vertically integrated cable providers.

## I. Introduction

The U.S. cable industry consists of two forms of oligopolies operating in close relation with each other in an estimated annual \$300 billion industry.<sup>1</sup> The first consists of multichannel video programming distributors (MVPDs) and the second consists of video programming providers. In the cable industry, MVPDs can vertically integrate with a video programming network to gain efficiency gains or achieve greater market share. This paper seeks to build upon prior works to understand if vertically integrated MVPDs have a tendency to carry their own affiliated program. Furthermore, the paper will use 2007 and 2010 industry data to investigate the possible existence of “quid pro quo” among vertically integrated MVPDs<sup>2</sup>. The existence of reciprocity is of interest for policy makers because major vertically integrated MVPDs can lockout or raise the cost of entry for new entrants such as Online TV by only sharing programming between themselves. Findings from this paper could have policy implications for vertical integrated video programming distributors in the cable industry as well as antitrust policy makers.

As mentioned above, the U.S. cable industry consists of two forms of oligopolies. The first consists of MVPDs such as cable television (CATV) systems, direct-broadcast satellite (DBS) providers, and wireline video providers. These companies usually derive revenue from monthly cable subscriptions, additional charges from premium channels, and rental fees from set-top boxes. Before the 1990s, CATV systems usually operated as the sole local cable supplier to a specific area; having received government sponsored monopolies and guaranteed returns.<sup>3</sup> These companies are generally known as Multiple System Operators (MSOs) and include firms such as Time Warner and Comcast. In the 1990s, new regulation and enforcement from the FCC allowed satellite operators to compete with the local cable

---

<sup>1</sup> Arango, T. (June 23, 2009). Cable TV's Big Worry: Taming the Web. *New York Times*.

<sup>2</sup> Note: Vertically integrated MVPDs that operate cable television are known as vertically integrated Multiple System Operators (MSOs). In this research paper, the author will use these two terms interchangeably.

<sup>3</sup> *Time Warner Entertainment Co., L.P. v. FCC*, 56 F.3d 151, 183 (DC Cir. 1995).

distributors.<sup>4</sup> Moreover, in this decade, telephone companies also entered the market; adding to the competition. However, even with increased competition, the cable industry still remains concentrated. Together, the top 10 MVPDs served 87 percent of subscribers in 2006.<sup>5</sup>

The second oligopoly consists of video programming networks that produce the content consumers watch. The programming network market is further divided into broadcasters and non-broadcasters. Broadcasters such as ABC, NBC, and CBS make their content available on cable TV and over-the-air. This is in contrast to non-broadcasters such as Viacom (owner of Comedy Central and MTV Networks) whose content is only available through cable subscriptions (Ammori 2010). The programming network market, like the programming distributors, is also a highly concentrated market dominated by a few dominant programming networks.<sup>6</sup> These companies mainly derive their revenue from advertising and per-subscriber fees.

From the above, it is apparent that the cable industry is a highly concentrated industry dominated by big distributors or programming networks that can leverage significant market power. In 2007 alone, out of the 565 national non-broadcast channels, 84 channels were affiliated with a cable operator; five of the top seven MVPDs held ownership interests in video programming networks.<sup>7</sup> Thus, the effects and impacts of vertical integration of MVPDs with video programming networks on consumers have long interested economists and antitrust policy makers.

There are two main hypotheses as to what MVPDs can gain from vertical integration. First, the efficiency hypothesis states that MVPDs merge with upstream programming networks to reduce double

---

<sup>4</sup> *Satellite Broadcasting and Communications Associate v. FCC*, 275 F.3d 337 (4<sup>th</sup> Cir. 2001).

<sup>5</sup> Ammori, M. (January 2010). TV Competition Nowhere: How the Cable Industry Is Colluding to Kill Online TV. *freepress*, 1-40.

<sup>6</sup> Cooper, M. and Turner, D. (Sept. 29, 2007). The Negative Effect of Concentration and Vertical Integration on Diversity and Quality in Video Entertainment. *35<sup>th</sup> Research Conference on Communication, Information and internet Policy (TPRC)*.

<sup>7</sup> Ammori, M. (January 2010). TV Competition Nowhere: How the Cable Industry Is Colluding to Kill Online TV. *freepress*, 1-40.

marginalization<sup>8</sup> and improve overall efficiency. Under this framework, price should fall and consumer welfare should increase because the MVPD becomes more efficient (Waterman and Weiss 1996). The other hypothesis is the vertical foreclosure hypothesis which is a focus of this research paper. In this framework MVPDs vertical integrate to achieve bigger market share or as close to monopoly power as possible to leverage market power to increase profitability or to hurt rival firms. In reality, vertical integration mergers may result from a combination of strategic foreclosure and efficiency incentives (Chipty 2001).

Under the foreclosure hypothesis, MVPDs can strategically foreclose on rivals by enforcing content lockout and violating net neutrality<sup>9</sup> (Ammori 2010). MVPDs can pressure video programmers to make it more expensive for rival MVPDs to carry their content or they can also choose not to carry rival affiliated programming to the detriment of consumers. Existing research has confirmed the existence of favoritism shown by MVPDs to their own affiliated programming networks (Waterman and Weiss 1996; Chipty 2001; Chen and Waterman 2007). Simply stated, MVPDs have a higher tendency to carry or promote their own programming over unaffiliated programming networks and rival affiliated MVPD programming networks. The second way that MVPDs can strategically foreclose recently is by violating net neutrality. Some MVPDs are involved in efforts to block or impair consumer's ability to access online videos (Ammiro 2010). The most recent and high-profile example is Comcast's degradation of high-definition online TV from providers such as ABC.com, Miro, and Vuze (Ammiro 2010).

As discussed above, favoritism shown by video programming distributors to affiliated video programming networks has been extensively studied (Waterman and Weiss 1996; Chipty 2001; Chen

---

<sup>8</sup> Double marginalization occurs when upstream and downstream firms have monopoly power and reduce outputs to the monopoly level. This results in two deadweight losses in the overall market. By having the firms vertically integrate, double marginalization can be eliminated, thus increasing consumer surplus. Ruth Maddigan and Janis Zaima (1985), "The Profitability of Vertical Integration," *Managerial and Decision Economics*, 178-179.

<sup>9</sup> Net Neutrality is the idea that all internet traffic should be treated equal. No matter who uploads or downloads data, or whatever the data involves, the networks should treat the transferred data packets on a first come, first serve basis. To do otherwise, would be data discrimination. Ulhs, A. (2007, April 19). Digital Divide: the Issue of Net Neutrality. *Imprint Magazine*.

and Waterman 2007). The preexisting research papers evaluate favoritism by mainly modeling carriage and subscription price. However, prior research papers focus exclusively on pre-2004 industry data prior to the recent growth of Online TV (Waterman and Weiss 1996; Ford and Jackson; 1997; Chipty 2001; Chen and Waterman 2007). With the introduction of Online TV, competition in video programming distribution has leveled the playing field by lowering the barrier of entry into the industry, thus increasing competition dramatically and prompting drastic actions from the incumbents of the industry (Ammori 2010). To the best of my knowledge, the possibility of existence of “quid pro quo” or collusion among MVPDs has not been extensively studied. As stated above, the purpose of this paper is to build upon existing research on favoritism shown by MVPDs to affiliated video programming networks. Furthermore, the paper will use 2007 and 2010 industry data to investigate the possible existence of “quid pro quo” among vertically integrated MVPDs. Findings from this paper could have policy implications for vertical integrated video programming distributors in the cable industry as well as antitrust policy makers.

The paper is organized into the following sections. Section 2 presents a review of the existing literature on vertically integrated video programming distributors and programmers. The theoretical framework of my research is presented in Section 3. Section 4 discusses the nature of the data used in this research paper. This is followed by Section 5 which presents the empirical methodology used in the analysis and the results of the empirical study. Lastly, Section 6 concludes and discusses policy implications.

## II. Literature Review

Vertical integration of video programming distribution and video programming networks has raised concerns that competition among the cable industry may be stifled and consumers harmed if cable companies strategically foreclose and deny programming to rival companies (Chipty 2001). As discussed below, the literature on vertical integration of the cable industry is primarily concerned with the effects of vertical integration on network carriage, subscriber number, and cable package price. Prior research uses structural models to model demand and supply curves for the cable industry to assess the effects of vertical integration on carriage and cable price. Using the results, the authors then assess the reasons behind the effects on the variables mentioned above and attribute the results to market foreclosure, efficiency gains, or even a combination of both. Five research studies are discussed in this section to focus on the effects of vertically integrated MVPDs specifically on consumer and producer welfare and favoritism shown by MVPDs to their own affiliated programming networks.

Waterman and Weiss (1996) are the first to investigate the existence of favoritism shown by MVPDs to their affiliated networks. Their study includes of 1,646 cable systems using 1987 and 1988 industry data. From the systems, they select eight major pay networks, some affiliated and some not, to study the effects of vertical integration. They model the cable industry's supply and demand functions and employ backward elimination and the Schwarz criterion to pick out significant system variables. They find that majority ownership ties to four major pay networks tended to carry their affiliated networks more frequently and rival networks less frequently than did the average nonintegrated system (Waterman and Weiss 1996). Furthermore, the authors conclude that vertical integration has little effect on prices, thus price is neutral. However, the study does admit that vertically integrated systems can still employ discounting strategies and other marketing techniques to favor affiliated over rival networks.

In Ford and Jackson (1997), the authors seek to address the effects of both horizontal and vertical integration of MSO on price and welfare. They assess the effects of ownership concentration and vertical integration by extending prior empirical research on cable television and prices. The paper follows earlier research by proposing a set of equations to model cable industry's supply and demand. The authors extend prior models by adding a programming cost function to the system in order to "account for the cost of the quality level chosen by the cable operator" (Ford and Jackson 1997). The set of equations are then used to assess the welfare effects increased by vertical integration and horizontal concentration. After running two-stage least squares estimates of the parameters proposed in their model, the authors find that vertical integration has a direct significant effect on lowering cable price (Ford and Jackson 1997). This is contrary to Waterman and Weiss (1996) where price is largely neutral. On the subject of consumer welfare, the paper finds that there is a small negative decrease of consumer welfare when MSOs vertically integrate due to the decrease in selection and quality of programming packages (Ford and Jackson 1997).

Another study that looks at the effects of vertical integration on network carriage, price, and consumer welfare is done by Chipty (2001). The study is more extensive than the two prior papers because it includes a more comprehensive list of MSOs and programming networks in its 1991 industry data. Chipty extends previous studies by examining the effects of vertically integrated ownership structure on program offerings. The study finds that vertical integration between cable operators and both premium and basic program services results in the exclusion of rival services. Furthermore, empirical evidence in the study shows that vertical integration results in efficiency gains. Therefore, contrary to Ford and Jackson (1997), Chipty shows that consumers in integrated markets are weakly better off than consumers in non-integrated markets because of these efficiency gains (Chipty 2001).

In addition, Chen and Waterman (2007) examine cable data from 2003 to study the effects of favoritism MSOs show to their own affiliated programming versus their rivals' programming. The paper is different from prior studies in that it examines the cable industry during the expansion of digital networks which were not present in the 1990s. In addition to reviewing prior studies on favoritism, the authors add another dimension to the Waterman and Weiss (1996) study by introducing another hypothesis. The authors introduce a positioning hypothesis which states that MSOs will place rival programming in more inaccessible tiers and affiliated programming in more accessible tiers. In this case, MSOs will place rival programming in the digital tier versus analog tier because consumers must pay more for the digital tier. The study looks at four distinct programming networks to test both the vertical integration and positioning hypotheses. The networks are chosen with the following criterion. First, the networks are chosen where there are both MSO affiliation and independent affiliation. Second, the networks are presumed to be close substitutes of each other. The four programming networks include outdoor entertainment, cartoon, basic movie service, and premium network. By assuming the standard error is normally distributed, the authors model the system using maximum likelihood probit with respect to carriage and analog tier.

From the results, the authors find that, except in a few cases, both vertical integration and positioning hypothesis hold (Chen and Waterman 2007). Specifically, Comcast and Time Warner favor their vertically integrated programming networks over their rivals and tend to place their rivals in more inaccessible tiers. In cases that MSOs were more likely to carry their rival programming, the MSOs placed rival programming with higher probability in the digital tier which is more inaccessible to consumers. The paper admits that the authors cannot "formally identify causes of the carriage or positioning differences [they] observe" or draw generalities from just four distinct programming networks and two MSOs. However, the pattern of relatively lower carriage of rival program suppliers remains a persistent phenomenon in the cable industry and requires more study. The paper provides

new insight into the cable industry and justifies that “cable industry cannot be understood only in terms of network carriage differentials. Tier positioning has opened a new dimension to the debate”.<sup>10</sup>

In the most recent study, Goolsbee (2007) takes a slightly different approach to document the evidence on vertical integration in the cable industry by examining television programming. The paper documents the prevalence of vertical integration and seeks to address whether integrated producers systematically discriminate against independent content in favor of their own content. The paper focuses on two types of video market – primetime broadcast programming and cable network carriage.

In broadcast programming, Goolsbee finds that integrated programmers have a lower standard (measured by advertisement revenues) for carrying their own shows versus independent shows created by the owners of rival broadcast networks (Goolsbee 2007). Furthermore, he finds that there is little evidence that vertically integrated networks enjoy major efficiencies in costs or revenues over their non-vertically integrated independent rivals (Goolsbee 2007). In all, he concludes that evidence on self-carriage rates show that vertically integrated cable systems are more likely to carry their own channels except in local regions where there is direct competition from DBS (Direct Broadcast Satellite). The finding is consistent with the four prior research papers detailed in this section.

For my research, I will first build upon existing research on favoritism shown by vertically integrated MVPDs. Also I examine the data for evidence of favoritism. Then I will extend the analysis to examine the possibility of “quid pro quo” among vertically integrated MVPDs. None of the prior research papers have examined this aspect of MVPD behaviors. My research will conclude with policy recommendations derived from my empirical analysis.

---

<sup>10</sup> D. Chen and D. Waterman. (2007). Vertical ownership, Program Network Carriage, and Tier Positioning in Cable Television: an Empirical Study, *Springer Science & Business Media B.V.* 2007, 250

### III. Theoretical Framework

In this research paper, I am investigating the effects of vertical integration on cable network carriage. Prior works mentioned in the last section examined the effects of vertical integration under the foreclosure hypothesis. The four research papers, Waterman and Weiss (1966), Ford and Jackson (1997), Chipty (2011), and Chen and Waterman (2007), all conclude that vertically integrated MVPDs are more likely to carry their own affiliated networks and are more likely to exclude their rivals. All four papers use similar approaches and methodology to evaluate the foreclosure hypothesis by modeling carriage supply and demand and running a regression controlling for demographics and headend systems on one year's worth of data.

In this paper, I will also examine a subset of MVPDs known as Multiple System Operators (MSOs) who own and operate cable television systems. I will see if MSOs tend to carry their own affiliated networks. Furthermore, I will explore a modified version of the foreclosure hypothesis. In this framework, I hypothesize that major vertically integrated MSOs will tend to carry each other's affiliated networks in a quid pro quo relationship. This relationship has not been explored by any existing literature. Under my hypothesis, major vertically integrated MSOs have an incentive to carry each other's affiliated networks. The incentive exists because vertically integrated MSOs in reciprocity relationships will directly benefit from an enlarged subscription base. Specifically, these vertically integrated MSOs will benefit from the per subscriber fee charged to other cable providers to carry their affiliated network. For example, if Time Warner carries Golf Channel (owned by Comcast) on all of its headend facilities, Comcast earns \$0.25<sup>11</sup> for each of Time Warner's subscribers per month and also increases Time Warner's cable program diversity and quality. Furthermore, advertisement on Comcast's affiliated networks will be worth more because of the increased viewership gained. This is in contrast to Time Warner carrying an

---

<sup>11</sup> Source: 2009: SNL Kagan Q4 2009 Multichannel Subscribers by DMA, March 29, 2010 (contains copyrighted and trade secret material distributed under license from SNL).

independent network, where additional viewership comes from districts that Time Warner already serves. This additional viewership is usually much smaller than a major vertically integrated MSO's subscribers base. Thus, vertically integrated MSOs are much more incentivized to carry other's affiliated networks. Moreover, this type of favoritism can lead to a lockout or a raise of cost of entry for new entrants into the cable industry.

To put it simply, the incentive structure is analogous to the iterated Prisoner's Dilemma from Game Theory. In the Prisoner's Dilemma, two prisoners are questioned by the police. If one testifies for the prosecution against the other (defects) and the other remains silent (cooperates), the defector gets a minor charge and the silent accomplice receives the full sentence. If both remain silent, both prisoners are set free. If each betrays the other, each receives a reduced sentence less than the full sentence but more than the minor charge. In essence, each prisoner must choose to betray the other or to remain silent. In the case of a single game, the Nash Equilibrium is where both prisoners choose to defect. However, in the iterated Prisoner's Dilemma, the incentive to defect can be overcome by the threat of punishment.

Similarly, in the cable industry, if Comcast and Time Warner do not carry each other's affiliated networks, both networks will suffer from fewer subscribers and from lower program diversity and quality. In the cases where one firm carries the other firm's affiliated networks and the other does not, only one of the firms will benefit greatly. However, in repeated games, one firm will retaliate by not carrying the other's channels. Lastly, if both Comcast and Time Warner carry each other's affiliated network, both vertically integrated MSOs will benefit from the additional subscribers accessible through the reciprocal relationship. Therefore, top vertically integrated MSOs have an incentive to carry each other's affiliated networks.

One may wonder if my hypothesis will conflict with the existing literature's conclusion that vertically integrated MSOs tend to exclude their rivals. After all, one may think that Time Warner and Comcast are in direct competition with each other and should thus be considered as rivals. However in reality, the big players of the cable industry rarely do compete with each other. They are usually municipal monopolies in local regions because of the high cost of entry. In fact these MSOs are usually in direct competition with Satellite and recently with online television. Furthermore, prior works did not conclude that the major vertically integrated MSOs excluded each other's networks. Instead, they only investigate and find exclusion among selected vertically integrated MSOs and independent channels but not among the top vertically integrated MSOs. Therefore, a quid pro quo relationship between the top vertically integrated MSOs can still exist under the foreclosure hypothesis.

For example, in Chipty 2001, the author explores 1919 cable systems operated by 340 cable system operators from 1991. The paper concludes that MSOs finds that vertical integration between cable operators and both premium and basic program services results in the exclusion of rival services. More specifically, premium movie operators are more likely to exclude basic movie services. In this paper, the author investigates substitute channels but do not explore all channels owned by a specific vertically integrated MSO. Also, in the most recent paper, Chen and Waterman 2007, the authors only examine 16 channels in four television categories with 2004 cable industry data. In their paper, they include Comcast and Time Warner in their regression, however, the networks investigated were owned by Time Warner and Comcast at different periods, not owned separately. Thus, the authors' findings do not necessarily show that Comcast and Time Warner exclude each other's affiliated networks.

In all, the theoretical framework used in this research paper does not necessarily contradict or refute past evidence of foreclosure hypothesis. Merely, this research paper will draw upon the rich array of past works on vertical integration in the cable industry to add to and extend the knowledge on vertical

integration's effects on network carriage for the major vertically integrated MSOs, specifically – Comcast, Time Warner, Cox, and Cablevision. Under my framework, I hypothesize that major vertically integrated MSOs will tend to carry each other's affiliated networks in a quid pro quo relationship.

## IV. Data

### A. TMS Data & SNL Kagan Data

The main data used in this research paper come from the Tribune Media Services (2007 & 2010 TMS Data). The TMS Data provides the channel lineups for every cable headend<sup>12</sup> in the US in February 2007 and 2010. An observation is a cable headend. There are roughly 8,000 headends in the U.S., which corresponds to around 656,858 duplicated observations in the dataset. For example, if headend 5343 carries n number of channels, the TMS data will have headend 5343 for n observations. For each headend, the data provides the geographic zip codes served by the headend, which is shown in Table 1. Furthermore, TMS data contains all of the channels offered at each headend and the channels' position within the offered bundle. This is designated by the variable Channel Position as listed in Table 1. In addition, the data set also provides the number of households that are serviceable by each headend and this is denoted by Sum of Household in the TMS data. Lastly, the data set also includes the MSO that maintains each of the headends that services a specific geographical location.

**Table 1: TMS Variable Definition**

<b>TMS Data Variable</b>	<b>Definition</b>
Headend Facility Identification Number	Each headend's unique Identification number.
Channel Position	Each channel at every headend is associated with a channel position or number within the offered bundle.
Channel Type	Each channel is associated with a channel type which can be one of the following - basic bundle, extended basic bundle, and digital basic bundle.
Sum of Household	This is the number of households each headend can serve within the geographic location
MSO Identification Number	For each cable headend, the data includes an unique MSO identification number that operates the respective cable headend

<sup>12</sup> A headend is a centralized facility that services and controls cable distribution in a geographic location in the U.S.

**Table 2: 2007 TMS Variable Summary**

<b>TMS Data Variable</b>	<b>Summary</b>	<b>Observations</b>	<b>Mean</b>	<b>Min, Max</b>
Headend Facility Identification Number	Each headend's unique Identification number with duplication	656858	12748.91	30, 26209
Channel Position	All of the channels offered at each headend and the channel position or number within the offered bundle	656858	185.2422	0, 9996
Channel Type	B: basic bundle EB: extended basic bundle DB: digital basic bundle	656858	2.1512	1,3
Sum of Household	This is the number of households each headend can serve within the geographic location	654051	35885.21	3, 1692332
MSO Identification Number	For each cable headend, the data includes an unique MSO identification number that operates the respective cable headend	656858	339.5985	2, 932

Doing a quick scan of the 2007 TMS data, I am able to come up with the total number of serviceable households in 2007 for each of the MSOs. These numbers are displayed below in Table 3. These numbers agree with existing sources<sup>13</sup> on the order of market share, with Comcast as the largest MSO with the most serviceable headend households followed by Time Warner, Cox, and Cablevision.

**Table 3: Total Serviceable Headend Households for MSOs in 2007**

<b>MSO</b>	<b>MSO's total serviceable headend households in 2007 (Millions)</b>
Cablevision	4.6
Comcast	40.4
Cox	7.9
Time Warner	23.7

<sup>13</sup> Top 25 Multichannel Video Programming Distributors, National Cable & Telecommunications Association, 2010. Marx, M. L. (2010). Economic Report on the Proposed Comcast-NBC Universal Transaction. Exhibit 3 of "Petition to Deny" filed by Patton Boggs LLP on behalf of Bloomberg L.P. in FCC Docket No. MB10-56, June 21, 2010.

The other data sets used in this research paper are SNL Kagan's Economics of Basic Cable Networks 2009 data and SNL Kagan TV Networks Peer Analysis 2010 which details the ownership of 175 cable networks as well as net advertising revenue. From the two datasets I chose the top 51 networks listed in Table 4 to investigate. These 51 networks provide a good representation of cable networks because they accounted for 87% of 2009's net advertisement revenue, which totaled at \$17,507.849 million. Out of the 51 networks, 15 networks are owned by the targeted four MSOs shown in Table 5. This accounts for approximately 29.4% of the networks considered in this paper. The other 36 networks considered are owned by major multi-network owners (defined as companies that own 5+ networks) with the exception of Bloomberg TV, Hallmark Channel, TV Guide Network, and WGN America.

**Table 4: 51 Networks**

ABC Family Channel	A&E	AMC	Animal Planet	BET
Bloomberg	Bravo	Cartoon Network	CMT	CNBC
CNBC World	CNN	Comedy Central	CSPAN	Discovery
Discovery Health Channel	Disney Channel	Disney XD	E! Entertainment Television	ESPN
ESPN2	Food Network	Fox News	Fox Network	Golf Channel
Hallmark Channel	HGTV	History Channel	Lifetime Television	Lifetime Movie Network
MSNBC	MTV	MTV2	Nickelodeon	Oxygen Network
Syfy	Speed Channel	Spike TV	TBS	TCM
TLC	TNT	Travel Channel	truTV	TV Guide Network
TV Land	USA	VH1	Weather Channel	WE tv
WGN America				

**Table 5: MSOs' Network Ownership**

<b>MSO</b>	<b>Programming Networks Owned by MSO</b>
Cablevision	AMC, WE Tv
Comcast	E! Entertainment Television, Golf Channel
Cox	Animal Planet, Discovery, Discovery Health, TLC, Travel Channel
Time Warner	Cartoon Network, CNN, TBS, TNT, TCM, truTV

## B. Alterations to the Data

Several alterations are made to the data before modeling and analysis are conducted. The first major alteration is to merge the TMS data sets from 2007 and 2010. By merging the two data sets, I am able to conduct my hypotheses across two time periods. The merging is accomplished by matching the community name and headend zip codes from the 2010 data set to the 2007 data set. This method is used because the unique headend identification number for headends changed from 2007 to 2010. The script to do the merging is included in Appendix B: Script 1 for review. After merging, the total households serviceable by all MSOs are seen in Table 6. This is due to the nature of the data set, since some headend facilities changed service areas during the three year period between 2007 and 2010.

**Table 6: Total Serviceable Headend Households for MSOs in Merged Data Set of 2007 and 2010**

<b>MSO</b>	<b>MSO's total serviceable headend households in 2007 (Millions)</b>
Cablevision	1.96
Comcast	19.5
Cox	5.24
Time Warner	19.9

After merging the 2007 and 2010 TMS data sets, further data alterations are made to come up with the data set to run multivariate OLS and probit regressions. For the Multivariate OLS regression, I found each of the 51 networks' percentage carriage on all tiers weighted by each headend's households for each MSO carrier facility. This is done by observing each headend to see if the headend carried Network A on any tier such as basic, extended basic, or digital basic. If the headend does carry Network A, then add the headend's serviceable households to the number of household receiving Network A carried by MSO A. Finally, the percentage carriage for Network A on MSO A is derived by dividing the number of households receiving Network A carried by MSO A by the total number of household serviceable by MSO A. The equation is shown below by Equation 1. This process is repeated for all MSOs – Cablevision, Comcast, Cox, and Time Warner. The script to this is attached in Appendix: B Script 2 and Script 3.

**Equation 1: Network Percentage Carriage**

$$\text{Network Percentage Carriage} = \frac{\text{Households Receiving Network A Serviceable to MSO B}}{\text{Total Households Serviceable by MSO B}}$$

Finally, I match network ownership for the 15 channels that are owned by the four major MSOs and create a data set that includes the following variables: (a) percentage carriage of a network on a MSO carrier, (b) year 2007, (c) year 2010, (d) MSO ownership, (e) 51 network dummy variables. The descriptions of each variable are described in Table 7. Overall, for each MSO carrier there will be 102 network observations for each of the 51 networks' percentage carriage.

**Table 7: Multivariate OLS Data Set Variable Description**

Variable	Description
Network Percentage Carriage on a MSO carrier	Value can range from 0-100. Network Percentage Carriage is the households receiving the network observation subscribed to the particular MSO carrier divided by the total number of households serviceable by the MSO carrier
Cablevision	1 if network is owned by Cablevision, 0 otherwise
Comcast	1 if network is owned by Comcast, 0 otherwise
Cox	1 if network is owned by Cox, 0 otherwise
Time Warner	1 if network is owned by Time Warner, 0 otherwise
Year 2007	1 if network observation comes from 2007, 0 otherwise
Year 2010	1 if network observation comes from 2010, 0 otherwise
Network Dummies listed in Table :51networks	1 if network observation comes from one of the 51 networks, 0 for the other 50 dummies. For example, if network observation comes from TNT, TNT = 1 and 0 for the rest of the 50 other network dummies

Table 8: Multivariate OLS Data Sample (listed below) details a sample of the data used for multivariate OLS. This particular sample comes from headends that are owned by Comcast. In the example below there are 6 observations which equates to 3 channels. The first two observations are the carriage percentage of Comcast's Households receiving CNN in 2007 and 2010. Notice that the network dummy variable for CNN is 1 for the first two observations and 0 for the rest. Also, notice that since CNN is

owned by Time Warner, the MSO Ownership dummy is 1 and 0 for the rest of the other MSO Ownership dummies. Similarly, the next two observations are Golf Channel's carriage percentage on Comcast's Household in 2007 and 2010. Lastly, the last two observations are Spike TV's carriage percentage on Comcast's Household in 2007 and 2010. Notice that since Spike TV is not owned by any of the four MSOs, hence the MSO ownership dummies are all 0.

**Table 8: Multivariate OLS Data Set Sample**

Carriage Percentage of Comcast's Households Receiving Network Observation	Year 2007	Year 2010	Cablevision	Comcast	Cox	Time Warner	CNN	Golf Channel	Spike TV
99.836	1	0	0	0	0	1	1	0	0
99.99	0	1	0	0	0	1	1	0	0
99.09	1	0	0	1	0	0	0	1	0
98.99	0	1	0	1	0	0	0	1	0
97.95	1	0	0	0	0	0	0	0	1
83.64	0	1	0	0	0	0	0	0	1
...									

The data set for the probit model is very similar to the one described above for the multivariate OLS regression. In the probit model data set, the network observations are derived from headends. Each headend will contribute 102 network observations (51 networks for 2 time periods). Also, Network Percentage Carriage on a MSO carrier is replaced by the variable, Carried. The variable, Carried, is 1 when the network observation for one of the 51 channels is carried. For example in Table 9 Probit Data Set, CNN is carried in both 2007 and 2010, so Carried = 1 for both observations. However, Spike TV is not carried in 2007 but is carried in 2010 so Carried = 0 in 2007 and Carried = 1 in 2010. The script to make the probit data is included in Appendix B: Script 5 for review.

**Table 9: Probit Data Set Sample**

Carried	Year 2007	Year 2010	Cablevision	Comcast	Cox	Time Warner	CNN	Golf Channel	Spike TV
1	1	0	0	0	0	1	1	0	0
1	0	1	0	0	0	1	1	0	0
1	1	0	0	1	0	0	0	1	0
1	0	1	0	1	0	0	0	1	0
0	1	0	0	0	0	0	0	0	1
1	0	1	0	0	0	0	0	0	1
...									

### C. Strengths & Limitations

TMS data is an appropriate data to use in this research paper for several reasons. First, the data contains all types of bundles or tiering for over 8000 headends in the United States. This is by far the largest headend facility sample as compared to datasets used in prior works. Furthermore, the TMS data offers different time period, allowing users to use fixed effects to control for factors unspecified in regression models.

There are several limitations that are apparent in the merged TMS 2007 & 2010 data. First, there are only half as many households in the merged data as compared to the 2007 TMS data. This limitation is simply due to the nature of the data. During the three year period, I postulate that many headends either merged with other headends or simply transitioned to digital. Therefore, it is impossible to match every headend from 2007 to 2010. It is important to note that there is no significant reason to believe that there is a systematic bias in the loss of data. In other words, there is no reason to suspect that the Comcast headends lost are headends that happen to not carry Time Warner affiliated channels. This will result in an upward bias of Time Warner affiliated networks carriage on Comcast facilities.

The second limitation of the data comes from some double counting of households in the data set. This occurs when two MSOs serve the same geographic area or zip code. For example, in the case that Time Warner and Comcast provide service in zip code 12345 which consists of 3000 households, in the

merged data, I counted an addition of 3000 for both Time Warner and Comcast for their respective total number of households. I believe that, this does not necessary skew my analysis since the number added is very insignificant.

## V. Empirical Methodology and Results:

### A. Multivariate OLS Regression

In order to determine the relationship between MSOs and their own affiliated networks and the relationship among the four MSOs, certain factors must be controlled for. In the section above, I determined the carriage percentage for each of the 51 networks received by each of the four MSOs' serviceable households in 2007 or 2010. These percentages are influenced by several factors. First, I expect that channels that are heavily demanded or have higher "quality" will have high carriage percentages. For example, ESPN is a highly demanded channel that will be carried in high percentage no matter who owns ESPN. Also, I expect that MSOs will be more likely to carry their affiliated networks and those of the other three MSOs. Thus, carriage percentages will be higher for these specific networks. In essence, a channel will have a higher carriage percentage because it is an affiliated network and or it is of superior quality. So in order to differentiate between the two forces, channel quality controls are introduced into the following regression in Equation 2:

#### Equation 2: Multivariate OLS Regression Equation

$$\begin{aligned} & \text{Network Percentage Carriage of 51 Networks on a MSO} \\ & = \beta_0 + \beta_1 \text{Cablevision} + \beta_2 \text{Comcast} + \beta_3 \text{Cox} + \beta_4 \text{TimeWarner} + \beta_5 \text{yr2007} \\ & + \beta_{6-56} \text{50 Network Dummy Variables} \end{aligned}$$

In the above multivariate OLS regression, the dependent variable is Network Carriage Percentage where it is defined in Equation 1. Cablevision is a dummy variable where it is 1 when the network observation is owned by Cablevision; Comcast is a dummy variable where it is 1 when the network observation is owned by Comcast; Cox is a dummy variable where it is 1 when the network observation is owned by Cox; TimeWarner is a dummy variable where it is 1 when the network observation is owned by Time Warner; yr2007 is a dummy variable where it is 1 when network observation is from 2007 and 0 when observation is from 2010; 50 network dummy variables is 1 when the network observation belongs to the respective network.

Prior works suggests that MSOs tend to carry their own affiliated network. Therefore, I expect the coefficients  $\beta_1$  to  $\beta_4$  to be positive when the networks are carried on their own facilities. For example,  $\beta_1$  will be positive when Cablevision is the facility carrier. Also, the theoretical framework of this analysis suggests that the coefficients  $\beta_1$  to  $\beta_4$  will be positive if reciprocity does exist among top MSO firms.  $\beta_5$  is expected to be negative because overall network percentage should increase over time due to technological advancements for headends. Lastly, the coefficient on the Network Dummy Variables is expected to be positive for channels that have high "quality." The regression results for Comcast, Time Warner, Cox, and Cablevision as carriers are shown in below in Table 10.

Table 10: Multivariate OLS Regression Results

	(1) Comcast	(2) Time Warner	(3) Cox	(4) Cablevision
Variable	Coefficients (%)	Coefficients (%)	Coefficients (%)	Coefficients (%)
Cablevision	30.78016*** (6.98911)	42.05964*** (5.78566)	31.01849*** (9.740735)	51.76011** (22.44481)
Comcast	40.08244*** (6.98911)	45.46025*** (5.78566)	47.40309*** (9.740735)	51.91953** (22.44481)
Cox	40.62694*** (6.98911)	46.61125*** (5.78566)	47.58477*** (9.740735)	53.1235** (22.44481)
Time warner	38.88855*** (6.98911)	45.05371*** (5.78566)	47.45132 *** (9.740735)	51.10869** (22.44481)
yr2007	-3.374823** (1.370677)	-2.418309** (1.13466)	-3.205545* (1.910315)	-12.03906*** (4.401789)
yr2010	(omitted)	(omitted)	(omitted)	(omitted)
ABC Family	30.3728*** (6.98911)	41.87142*** (5.78566)	31.03939*** (9.740735)	53.35168** (22.44481)
AMC	(omitted)	(omitted)	(omitted)	(omitted)
A&E	30.82817*** (6.98911)	42.17659*** (5.78566)	41.78876*** (9.740735)	53.31712** (22.44481)
Animal Planet	-4.261906 (6.98911)	-20.29932*** (5.78566)	-16.3747* (9.740735)	-52.32391 (22.44481)
BET	22.03884*** (6.98911)	35.22475*** (5.78566)	30.74042*** (9.740735)	52.13113** (22.44481)
Bloomberg	17.16102** (6.98911)	46.18912*** (5.78566)	42.19633*** (9.740735)	10.45435 (22.44481)
Bravo	26.57013*** (6.98911)	34.01431*** (5.78566)	30.98741*** (9.740735)	50.55022** (22.44481)
CMT	37.70736*** (6.98911)	46.60379*** (5.78566)	47.33634*** (9.740735)	12.22781 (22.44481)
CNBC	39.9805*** (6.98911)	45.32888*** (5.78566)	47.45132*** (9.740735)	52.06849** (22.44481)
CNC World	-48.00352*** (6.98911)	35.77376*** (5.78566)	-52.39807*** (9.740735)	-46.64832** (22.44481)
CNN	1.923466 (6.98911)	2.680515 (5.78566)	.1506119 (9.740735)	2.242996 (22.44481)
CSPAN	40.46968*** (6.98911)	46.37805*** (5.78566)	47.5537*** (9.740735)	52.62027** (22.44481)
Cartoon Network	.1630783 (6.98911)	.7221947 (5.78566)	.0852203 (9.740735)	1.305496 (22.44481)
Comedy Central	38.60139*** (6.98911)	45.51491*** (5.78566)	47.40309*** (9.740735)	52.63972** (22.44481)
Discovery	.1555748 (6.98911)	-.4032593 (5.78566)	.0171585 (9.740735)	.2281876 (22.44481)
Discovery Health	-2.288731 (6.98911)	.2280159 (5.78566)	-.0235443 (9.740735)	-98.50526*** (22.44481)
Disney Channel	40.12402*** (6.98911)	46.53072*** (5.78566)	47.43514*** (9.740735)	53.08669** (22.44481)
Disney XD	38.25156*** (6.98911)	42.2052*** (5.78566)	41.93386*** (9.740735)	8.965306 (22.44481)
E! Entertainment	(omitted)	(omitted)	(omitted)	(omitted)
ESPN	40.88233*** (6.98911)	47.10832*** (5.78566)	45.28505*** (9.740735)	53.35168** (22.44481)
ESPN2	40.63982*** (6.98911)	46.95688*** (5.78566)	39.40418*** (9.740735)	53.31524** (22.44481)
Fox Network	40.16631*** (6.98911)	46.22481*** (5.78566)	47.41967*** (9.740735)	18.07403 (22.44481)
Food Network	39.80392*** (6.98911)	45.85418*** (5.78566)	47.41967*** (9.740735)	17.57596 (22.44481)
Fox News	40.08705***	46.27943***	47.53654***	53.11099**

	(6.98911)	(5.78566)	(9.740735)	(22.44481)
<b>Golf channel</b>	-1.1461639 (6.98911)	-1.230782 (5.78566)	-.9502602 (9.740735)	-39.78787* (22.44481)
<b>HGTV</b>	40.50625*** (6.98911)	46.01719*** (5.78566)	47.57781*** (9.740735)	-15.35486 (22.44481)
<b>Hallmark</b>	38.316*** (6.98911)	43.63978*** (5.78566)	47.26202*** (9.740735)	12.92433 (22.44481)
<b>History Channel</b>	40.41605*** (6.98911)	44.30084*** (5.78566)	46.46941*** (9.740735)	52.83069** (22.44481)
<b>Lifetime</b>	40.74496*** (6.98911)	46.86864*** (5.78566)	47.54038*** (9.740735)	53.27499** (22.44481)
<b>Lifetime Movies</b>	35.86951*** (6.98911)	42.50425*** (5.78566)	41.23562*** (9.740735)	-45.38176** (22.44481)
<b>MSNBC</b>	37.46795*** (6.98911)	45.85925*** (5.78566)	47.46928*** (9.740735)	51.05355** (22.44481)
<b>MTV</b>	40.60443*** (6.98911)	45.87247*** (5.78566)	42.7984*** (9.740735)	52.28503** (22.44481)
<b>MTV2</b>	38.1658*** (6.98911)	45.14111*** (5.78566)	45.3719*** (9.740735)	51.10869** (22.44481)
<b>Nickelodeon</b>	40.86177*** (6.98911)	47.32674*** (5.78566)	47.60193*** (9.740735)	53.35168** (22.44481)
<b>Oxygen</b>	12.51905* (6.98911)	39.19732*** (5.78566)	30.81118*** (9.740735)	10.22104 (22.44481)
<b>Speed Channel</b>	38.71402*** (6.98911)	45.25093*** (5.78566)	47.44379*** (9.740735)	46.84697** (22.44481)
<b>SpikeTV</b>	40.74413*** (6.98911)	47.73422*** (5.78566)	47.60193*** (9.740735)	53.35168** (22.44481)
<b>Syfy</b>	38.71718*** (6.98911)	46.73814*** (5.78566)	47.60193*** (9.740735)	52.49116** (22.44481)
<b>TBS</b>	1.619637 (6.98911)	1.268333 (5.78566)	.1506119 (9.740735)	2.242996 (22.44481)
<b>TCM</b>	-3.868309 (6.98911)	-4.005062 (5.78566)	-10.55888 (9.740735)	-5.175224 (22.44481)
<b>TLC</b>	(omitted)	(omitted)	(omitted)	(omitted)
<b>TNT</b>	1.941216 (6.98911)	2.541004 (5.78566)	.1506119 (9.740735)	2.242996 (22.44481)
<b>TV Guide</b>	18.99206*** (6.98911)	17.40536*** (5.78566)	37.10823*** (9.740735)	-44.67368*** (22.44481)
<b>TVland</b>	31.68995*** (6.98911)	40.4297*** (5.78566)	30.85339*** (9.740735)	31.20866** (22.44481)
<b>Travel Channel</b>	-4.308399 (6.98911)	-.6726646 (5.78566)	-.1816864 (9.740735)	-13.28088 (22.44481)
<b>truTV</b>	(omitted)	(omitted)	(omitted)	(omitted)
<b>USA</b>	40.89282 (6.98911)	47.10832*** (5.78566)	47.60193*** (9.740735)	53.35168** (22.44481)
<b>VH1</b>	29.91555 (6.98911)	41.05941*** (5.78566)	30.63519*** (9.740735)	52.31489** (22.44481)
<b>WE tv</b>	-6.841473 (6.98911)	-2.395538 (5.78566)	-26.71465*** (9.740735)	.2090912 (22.44481)
<b>WGN America</b>	1.578417 (6.98911)	22.15636*** (5.78566)	42.63261*** (9.740735)	-40.63986* (22.44481)
<b>Weather</b>	40.6686*** (6.98911)	46.731*** (5.78566)	47.46791*** (9.740735)	53.18501** (22.44481)
<b>Constant</b>	60.78871*** (4.98934)	53.47493*** (4.130229)	54.00084*** (6.953653)	52.66785*** (16.02275)
Observations	104	104	104	104
R-squared	0.9028	0.8198	0.8524	0.8014
R-squared Adjusted	0.8037	0.6360	0.7018	0.5988

Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

In column (1) of Table 10, the coefficient for Comcast is very significant ( $p < 0.01$ ) and positive with a value of 40.08. This suggests that Comcast strongly favors its own affiliated networks, Golf Channel and E! Entertainment. The evidence suggests that if a channel is owned by Comcast, its network percentage on Comcast facilities will get an additional carriage of 40.08%. The result is similar to prior works cited in Section 2. Also, the coefficients on Cablevision, Cox, and Time Warner are positive and are also very significant ( $p < 0.01$ ). The results suggest that Comcast also favors channels affiliated with the other three vertically integrated MSOs strongly. A channel owned by Time Warner will have a higher percentage carriage by 38.89% just because it is owned by Time Warner. This is similar for channels owned by Cox and Cablevision where the channels all have a higher network carriage percentage of 40.63% and 30.78% respectively. These coefficients are all reported with a significance of  $p < 0.01$ .

The coefficient on year 2007 variable is slightly negative, -3.37%, but very significant ( $p < 0.01$ ). This is as expected since network carriage tends to increase over time as more and more headends are upgraded and are not capacity constrained. In addition, many of the network dummy variables are also very significant. For example, if the channel is ESPN, Comcast will carry it with an additional network carriage of approximately 41%. Overall, 35 of the 51 channels are significant and indicate that the variables are well chosen and control for quality well.

As one may notice from the regression results, several of the network dummies are omitted from the regression. These variables are omitted because of collinearity. For example, if E! Entertainment Channel and Golf Channel are almost always offered together, one of the network dummy will be omitted due to collinearity. In this case, E! Entertainment is excluded from the regression. In essence, the added observations do not add more information to the regression and are therefore omitted. Lastly, the R-squared of the regression is 0.9028 and the adjusted R-squared is 0.8037. This suggests that the regression on Comcast as the carrier explains around 80% of the variation observed in the data.

Similar results are also found in column (2) where Time Warner is the carrier. Just like Comcast, Time Warner also has a tendency to carry its own affiliated networks with higher carriage of 45.03% ( $p < 0.01$ ). Furthermore, Time Warner also strongly favors networks owned by Comcast, Cox, and Cablevision. The coefficients are positive, economically significant, and statistically significant ( $p < 0.01$ ). Furthermore, 38 of the 51 channel dummies are statistically significant. Similar results are also found in column (3) and column (4) where the carriers are Cox and Cablevision respectively.

From the results above, it is apparent that all four MSOs have similar results. All vertically integrated MSOs have a strong tendency to carry their own respective affiliated networks with a high statistical significance. This is what is expected from prior works and in leading theories. More importantly, the above results suggest that reciprocity does indeed exist among the four big vertically integrated MSOs. The coefficients for all four MSOs for each MSO as the carrier are positive, economically significant, and statistically significant. For example, when Comcast is the carrier, the Time Warner coefficient is positive and when Time Warner is the carrier, the Comcast coefficient is also positive. Therefore, the evidence supports the conclusion that a reciprocity relationship of quid pro quo does exist between Comcast and Time Warner. In all, the data suggest the reciprocity relationship exists for all of the possible pairs of MSOs pairs.

After completing the multivariate OLS Regression, I now have enough information to calculate the additional revenue that each vertically integrated MVPD cable provider directly gains in the reciprocity relationship. This involves a simple calculation that is outlined below in Equation 3. Equation 3 uses several variables. The first variable, MSO A Carrier's Subscribers, details the number of subscribers in millions of households for each of the vertically integrated MVPD cable provider explored in this research paper. The values for MSO B's Network Subscription Cost per Month per Subscriber can be found in Table 12. This variable accounts for how much any MVPD must pay the vertically integrated

MVPD cable provider to carry its programming networks. Lastly, the Network Percentage Carriage of MSO B's Affiliated Networks is found from regression results from Table 10 (listed above).

### Equation 3: Additional Subscriber Revenue Equation

*Annual Additional Revenue for MSO B*

$$= \sum_{i=0}^n 12 \text{ months}$$

- \* *MSO A's Network Subscription Cost per Month per Subscriber*
- \* *MSO<sub>i</sub> Carrier's Subscribers*
- \* *Network Percentage Carriage of MSO A's Affiliated Networks on MSO<sub>i</sub>*

**Table 11: 2009 Market Shares for MSOs**

MSO Name	Subscribers (millions of households)
Comcast	23.78
Time Warner Cable	12.98
Cox	5.32
Cablevision	3.20

Source: 2009: SNL Kagan Q4 2009 Multichannel Subscribers by DMA, March 29, 2010 (contains copyrighted and trade secret material distributed under license from SNL).

**Table 12: 2009 Cable Networks and Subscriber Fees**

Programming Network Name	Vertically Integrated MSO Owner	Subscription Cost per Month (\$/Month per Subscriber)
AMC	Cablevision	0.23
WE tv	Cablevision	0.11
E! Entertainment Television	Comcast	0.20
Golf Channel	Comcast	0.25
Animal Planet	Cox	0.07
Discovery	Cox	0.25
Discovery Health	Cox	0.12
TLC	Cox	0.16
Travel Channel	Cox	0.08
Cartoon Network	Time Warner	0.17
CNN	Time Warner	0.51
TBS	Time Warner	0.49
TNT	Time Warner	0.99
TCM	Time Warner	0.26
truTV	Time Warner	0.10

Source: 2009: SNL Kagan Q4 2009 Multichannel Subscribers by DMA, March 29, 2010 (contains copyrighted and trade secret material distributed under license from SNL).

The additional revenue that Comcast will receive from Time Warner carrying Comcast's affiliated networks can be calculated in the following manner using Equation 3 listed above. Looking at Table 12, Entertainment Television's and Golf Channel's subscription costs per month per subscriber are \$0.20 and \$0.25 respectively. From the regression results, Time Warner carries Comcast affiliated channels with an addition of 45.46% in network carriage. This equates to 5.9 million households of Time Warner's 12.98 million households subscriber base. Substituting these values into Equation 3 results in Comcast earning an additional \$31.864 million per year. This is assuming no further subscriber growth.

Also note that this is just the direct monetary gain that Comcast will experience, without accounting for the increase of advertisement value of gaining another 5.9 million households in viewership. Similarly for Cox and Cablevision, Comcast will earn an additional \$13.62 million per year and \$8.66 million per year respectively. Overall Comcast will earn \$54.14 million per year and will gain 10.03 million households in viewership for its programming networks. Similar calculations can be made for Time Warner, Cox, and Cablevision. In all, Time Warner will earn \$403.745 million with a gain of 13.35 million households in viewership. Cox will earn \$141.598 million with 17.35 million households more in viewership. Lastly Cablevision receives \$58.87 in programming network subscriptions and a gain of 14.43 million households in viewership.

Next, the gain of additional viewership from the reciprocity relationships allows me to calculate the percentage change of advertisement revenue. Keith Brown and Roberto Cavazos (2005) cite that a 1% increase in the share of a broadcast network program is associated with a 0.39% increase in advertising price per viewer. I find the percentage change of advertisement revenue by Equation 4 shown below.

The final calculations are shown below in Table 13.

**Equation 4: Percentage Change in Advertisement Revenue Equation**

$$\% \Delta \text{ in Advertisement Revenue} = (1 + \% \Delta \text{ viewers}) * (1 + e_p * \% \Delta \text{ viewers}) - 1$$

$\% \Delta \text{ viewers}$ : the percentage change of viewers

$e_p$ : the price elasticity for advertisement revenue

**Table 13: MSO Benefits**

MSO	Subscription Revenue (million)	Viewership Gain (million of households)	%Δ in Advertisement Revenue
Cablevision	\$59	14.43	20.9
Comcast	\$54	10.03	14.3
Cox	\$142	17.35	25.3
Time Warner	\$404	13.35	19.3

From Table 13, Cablevision can expect up to an increase of 20.9% in percentage change in advertisement revenue if it carries affiliated networks owned by Comcast, Cox, and Time Warner.

Similarly, Comcast, Cox, and Time Warner can expect up to 14.3%, 25.3%, and 19.3% increase in advertisement revenue while carrying other MSOs' affiliated networks in a quid pro quo relationship.

These numbers point to the economic significance of the reciprocal relationship. Therefore, it is easy to see why major vertically integrated MSOs have an incentive to participate in a quid pro quo relationship.

**B. Probit Regression**

Another approach to model the relationship between vertically integrated MSOs and their affiliated networks is to use the probit regression. The probit regression differs from the multivariate OLS regression in the sense that it models for the probability for any channel to be carried on a certain MSO. Furthermore, probit regression has an advantage of bounding the probability of network carriage between 0 and 1 while this is not the case for the multivariate OLS. The same variables are controlled for in the probit regression as in the multivariate OLS regression. In all, the probit regression is constructed the same way as the multivariate OLS regression and the probit regression equation is listed in Equation

5.

**Equation 5: Probit Regression Equation**

$$\begin{aligned}
 &Pr(\text{Carried on MSO A Carrier} = 1 | \text{MSO Variables, yr2007, 50 Network Dummy Variables}) \\
 &= \Phi(\beta_0 + \beta_1 \text{Cablevision} + \beta_2 \text{Comcast} + \beta_3 \text{Cox} + \beta_4 \text{TimeWarner} + \beta_5 \text{yr2007} \\
 &+ \beta_{6-56} \text{50 Network Dummy Variables})
 \end{aligned}$$

In the above probit regression, the dependent variable is “Carried on MSO A carrier” where it is 1 whenever any of the 50 Network Dummy Variables is 1 and 0 otherwise. Cablevision is a dummy variable where it is 1 when the network observation is owned by Cablevision; Comcast is a dummy variable where it is 1 when the network observation is owned by Comcast; Cox is a dummy variable where it is 1 when the network observation is owned by Cox; TimeWarner is a dummy variable where it is 1 when the network observation is owned by Time Warner; yr2007 is a dummy variable where it is 1 when network observation is from 2007 and 0 when observation is from 2010; 50 network dummy variables is 1 when the network observation belongs to the respective network. The marginal effects of the probit regression are displayed below in Table 14.

Table 14: Probit Regression Results

	(1) Comcast	(2) Time Warner	(3) Cox	(4) Cablevision
Variable	Marginal Effects	Marginal Effects	Marginal Effects	Marginal Effects
Cablevision	-.0176832*** (.00003)	-.0095441*** (.00005)	-.0652727*** (.00009)	-.013583*** (.00029)
Comcast	.0204101*** (.00001)	.0152819*** (.00002)	.0155807*** (.00001)	-.6898795*** (.00064)
Cox	.0272614*** (.00001)	.0286742*** (.00002)	.0249759*** (.00003)	.0384459*** (.00015)
Time warner	.0214594*** (.00001)	.0196399*** (.00002)	.019292*** (.00002)	-.0444251*** (.00036)
yr2007	-.0226956*** (.00001)	-.0214536*** (.00001)	-.0153906*** (.00001)	-.1821016*** (.0001)
yr2010	(omitted)	(omitted)	(omitted)	(omitted)
ABC Family	.0012645*** (.00002)	.0041045*** (.00004)	.0005607*** (.00002)	(omitted)
AMC	.0113353*** (.00001)	.0114234*** (.00003)	.0129897*** (.00001)	-.0065813*** (.00025)
A&E	.0023712*** (.00002)	.0056942*** (.00003)	.0116155*** (.00001)	.0464189*** (.00006)
Animal Planet	-.1393975*** (.00025)	-.2969389*** (.0003)	-.6243482*** (.00223)	-.0625357*** (.00071)
BET	-.0248784*** (.00003)	-.032139*** (.00007)	.0001173*** (.00002)	-.0073286*** (.00028)
Bloomberg	-.0421068*** (.00004)	.0232714*** (.00001)	.0118853*** (.00001)	-.7236912*** (.00061)
Bravo	-.0095252*** (.00002)	-.0377593*** (.00007)	.0004857*** (.00002)	-.065192*** (.00042)
CMT	.0160157*** (.00001)	.0250646*** (.00001)	.0150367*** (.00001)	-.7024876***
CNBC	-.6388202*** (.00015)	-.0287565*** (.00006)	(omitted)	(omitted)
CNC World	.0194842*** (.00001)	.0194409*** (.00002)	.0152093*** (.00001)	-.0097887*** (.00028)
CNN	.0194609*** (.00001)	(omitted)	(omitted)	(omitted)
CSPAN	.0201878*** (.00001)	.0241011*** (.00001)	.0155353*** (.00001)	.0126129*** (.00021)
Cartoon Network	.0021132*** (.00003)	.007413*** (.00004)	.0067286*** (.00008)	.0278966*** (.00011)
Comedy Central	.0175181*** (.00001)	.0203416*** (.00002)	.015122*** (.0001)	.0134384*** (.00021)
Discovery	-.0871941*** (.0002)	.005078*** (.00006)	-.01105*** (.00042)	-.9667306*** (.00003)
Discovery Health	.0102153*** (.00004)	-.0069855*** (.00008)	(omitted)	(omitted)
Disney Channel	.0196634*** (.00001)	.0248453*** (.00001)	.0151969*** (.00001)	.0335818*** (.00013)
Disney XD	.0169025*** (.00001)	.0038165*** (.00004)	.0111488*** (.00001)	-.7406644*** (.00057)
E! Entertainment	.0029201*** (.00004)	.0093778*** (.00004)	.01136*** (.00002)	.0510874*** (.00004)
ESPN	.0205072*** (.00001)	.0265*** (.00001)	.0095031*** (.00001)	.0462766*** (.00006)
ESPN2	.0214634*** (.00001)	.0271624*** (.00001)	.0135779*** (.00001)	(omitted)
Fox Network	.0197402*** (.00001)	.0234278*** (.00001)	.0151498*** (.00001)	-.6306037*** (.00071)
Food Network	.0192303***	.0217711***	.0151498***	-.6383357***

	(.00001)	(.00002)	(.00001)	(.00071)
<b>Fox News</b>	.0196279*** (.00001)	.0236261*** (.00001)	.0154518*** (.00001)	.034778*** (.00012)
<b>Golf channel</b>	(omitted)	(omitted)	(omitted)	***
<b>HGTV</b>	.0202674*** (.00001)	.0226215*** (.00002)	.0157047*** (.00001)	-.8989126*** (.00024)
<b>Hallmark</b>	.0170001*** (.00001)	.0120147*** (.00003)	.0149514*** (.00001)	-.6943483*** (.00065)
<b>History Channel</b>	.0201258*** (.00001)	.0153704*** (.00002)	.0144007*** (.00001)	.0217265*** (.00018)
<b>Lifetime</b>	.0124916*** (.00001)	.0054779*** (.00003)	.0105535*** (.00001)	-.9654776*** (.00003)
<b>Lifetime Movies</b>	.0207622*** (.00001)	.0261624*** (.00001)	.0154846*** (.00001)	.0445087*** (.00007)
<b>MSNBC</b>	.0155276*** (.00001)	.0217895*** (.00002)	.0152478*** (.00001)	-.0433148*** (.00037)
<b>MTV</b>	.016831*** (.00001)	.018737*** (.00002)	.0137935*** (.00001)	-.0456175*** (.00038)
<b>MTV2</b>	.0204431*** (.00001)	.0218651*** (.00002)	.0122971*** (.00001)	-.0012028*** (.00026)
<b>Nickelodeon</b>	.0212311*** (.00001)	.0281168*** (.00001)	(omitted)	(omitted)
<b>Oxygen</b>	-.0658832*** (.00006)	-.0124478*** (.00005)	-.0021959*** (.00003)	-.7260128*** (.0006)
<b>Speed Channel</b>	.0176863*** (.00001)	.0191697*** (.00002)	.0151914*** (.00001)	-.1450156*** (.00057)
<b>SpikeTV</b>	.0207514*** (.00001)	(omitted)	(omitted)	(omitted)
<b>Syfy</b>	.0177348*** (.00001)	.0256097*** (.00001)	(omitted)	.0072054*** (.00023)
<b>TBS</b>	.0163101*** (.00001)	.014018*** (.00003)	(omitted)	(omitted)
<b>TCM</b>	-.037861*** (.00006)	-.0391699*** (.00009)	-.2945887*** (.00085)	-.0740405*** (.00035)
<b>TLC</b>	(omitted)	(omitted)	(omitted)	(omitted)
<b>TNT</b>	.0196508*** (.00001)	.0277855*** (.00001)	(omitted)	(omitted)
<b>TV Guide</b>	-.0375178*** (.00004)	-.1518396*** (.00014)	.0064553*** (.00002)	-.9648873*** (.00003)
<b>TVland</b>	.0016552*** (.00002)	-.0058817*** (.00004)	-.0021644*** (.00003)	-.4366885*** (.0008)
<b>Travel Channel</b>	-.1452793*** (.00026)	-.0141809*** (.00009)	-.0521745*** (.00081)	-.5186696*** (.00132)
<b>truTV</b>	(omitted)	(omitted)	(omitted)	(omitted)
<b>USA</b>	.0217493*** (.00001)	.0271624*** (.00001)	(omitted)	(omitted)
<b>VH1</b>	(omitted)	(omitted)	(omitted)	(omitted)
<b>WEtv</b>	(omitted)	(omitted)	(omitted)	(omitted)
<b>WGN America</b>	-.1212673*** (.00008)	-.1180377*** (.00012)	.0120199*** (.00001)	-.9613684*** (.00004)
<b>Weather</b>	.0205767*** (.00001)	.025605*** (.00001)	.0152447*** (.00001)	.0385514*** (.00011)
<b>Constant</b>	0.98082768	0.97218328	0.98652591	0.95439598
<b>Observations</b>	1991698104	984363058	440608812	160633080
<b>Pseudo R<sup>2</sup></b>	0.3768	0.1692	0.2681	0.6648

Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

In the first column of Table 14, the coefficient for Comcast has a value of .0204 and is statistically significant ( $p < 0.01$ ). The regression results suggest that Comcast slightly favors its own affiliated networks with a 2.04% increase of network carriage. In short, this means that if a channel is owned by Comcast, the network has an additional 2.04% chance that will be carried on Comcast's facilities. This result still agrees with the foreclosure hypothesis. Similar results are found for Cox and Time Warner where the coefficients are 2.73% and 2.15%. The surprising result comes from Cablevision where the coefficient is -1.77% and is also statistically significant ( $p < 0.01$ ). This suggests that Comcast tends to carry less of Cablevision's affiliated networks. The result is in contrast to the result from the multivariate OLS result, where the Cablevision coefficient is positive.

The coefficient on year 2007 variable is slightly negative, -2.27%, but very significant ( $p < 0.01$ ). This is as expected since network carriage tends to increase over time, as headends are upgraded and are not capacity constrained. In addition, all of the 46 network dummy variables are significant, which indicates that the variables are well chosen. Furthermore, the network dummy variables that are omitted because of collinearity for the same reasons as explained above in the multivariate OLS regression results. Lastly, the pseudo  $R^2$  for column (1) is .3768. This means that the regression results only explain approximately 37.68% of the variations in the data.

Similar probit results are also found for Time Warner shown in column (2). Just like Comcast, Time Warner also has a slight tendency to carry its own affiliated networks with an addition of 1.96% in carriage probability ( $p < 0.01$ ). Furthermore, Time Warner also slightly favors networks owned by Comcast and Cox. The coefficients are positive and statistically significant ( $p < 0.01$ ). Furthermore, 43 of the 50 channel dummies are statistically significant. In this probit regression, some variables are omitted again because of collinearity. In addition, CNN and Spike TV are omitted because these variables predict

carriage success one hundred percent and therefore do not add additional information to the regression. Similar results are also found in column (3) carrier is Cox respectively.

Some surprising results are found in column (4) where Cablevision is the carrier. The first surprising result is that Cablevision is slightly less likely to carry its own affiliated networks by 1.36%. Looking at existing literature, this is rather a surprising result. Existing literature all agree that vertically integrated MVPDs tend to carry more of its affiliated networks. Furthermore, Cablevision is less likely to carry Comcast affiliated networks by 68.99% and Time Warner affiliated networks by 4.44%. These coefficients are all statistically significant.

The probit regression results provide a mixed conclusion. Three of the four vertically integrated MSOs have similar results in that the MSOs are slightly more likely to carry its own affiliated networks and carry the other MSOs except Cablevision. However, Cablevision's probit regression results paint a very different perspective. The evidence suggests that Cablevision slightly favors Cox's networks, favors less of other MSOs' affiliated networks, and less of its own affiliated networks. This is a different conclusion in contrast to existing literature and multivariate OLS regression results.

Given the discrepancies between the probit regression results and the multivariate OLS results, I would weigh the probit regression results less for several reasons. First, there are several channels that are omitted in the probit regression, because those channels predict success one hundred percent of the time. These channels are left out of the regression because they do not add additional information for the probit regression. However, these channels are important in predicting network carriage, and so it is unsurprising that the pseudo  $R^2$  for the probit regressions are generally lower. Second, every network in the probit regression is modeled as carried equally, regardless of what network it is. This is in contrast to multivariate OLS, where network carriage is modeled specifically for each network which can account for more individual variations in network carriage. Hence, this is reflected upon in the lower values of

the pseudo  $R^2$  for the probit regression results. In light of these reasons, I would weigh the probit regression results less than the multivariate OLS regression results.

## VI. Conclusion

Existing literature has shown that vertically integrated MVPDs are more likely to carry their own affiliated networks and less likely to carry rivals to their affiliated networks. In the existing studies, the authors use models to look primarily at substitute networks owned by non-vertically integrated MVPDs. In this paper, however, I investigate the effects of vertical integration using 2007 and 2010 data on a subset of MVPDs known as vertically integrated MSOs. Following the existing literature's thought on the foreclosure hypothesis, I test whether vertically integrated MSOs have an increased tendency to carry their own affiliated networks. In addition, I explore a new dimension of the effects of vertical integration in the cable industry. I look at network carriage behavior of four major vertically integrated MSOs: Comcast, Cablevision, Cox, and Time Warner. I postulate that vertically integrated MSOs have an incentive to carry each other's affiliated networks in a quid pro quo relationship.

Two approaches are taken to evaluate the hypotheses mentioned above. The first approach is to construct a multivariate OLS regression, and the second approach is to construct a probit regression. The multivariate OLS results are consistent with prior findings, statistically, and economically significant. The evidence suggests that vertically integrated MSOs have a strong tendency to carry their own respective affiliated networks. Furthermore, the data suggest that reciprocity relationships exist for all of the possible pairs of MSOs (Comcast, Cablevision, Cox, and Time Warner) examined in the analysis. While the probit regression seems appropriate to model the probability of network carriage, the probit regression results suggest mix conclusions. Therefore, as mentioned above, there are several reasons why the probit regression results are weighed less in the analysis.

The findings in this paper add value to research community in several ways. First, the results from the multivariate OLS regression are consistent with existing literature's results. The regression shows that even in an era of rapid technology changes and changes in the market share with 2007 and 2010 data,

vertically integrated MSOs still have a tendency to carry their own affiliated networks. Second, the paper furthers our knowledge of network carriage behaviors of major vertically integrated MVPD cable providers. The paper outlines several incentive structures that can attract major vertically integrated MVPD cable providers to participate in quid quo pro relationships.

In addition, the evidence presented in the paper can have policy implications for vertical integrated video programming distributors in the cable industry as well as antitrust policy makers. The existence of reciprocity is of interest for policy makers because major vertically integrated MVPDs can lockout or raise the cost of entry for new entrants by only sharing programming among themselves. Of course, further studies on welfare effects of reciprocity relationships of major vertically integrated MVPD cable providers will be needed to understand if the general public benefits from these relationships. However, the paper contributes to our understanding the relationships among major vertically integrated MVPD cable providers. It is an additional perspective on the effects of vertical integration on MVPD cable providers in an ever changing industry and provides ample extensions for future researchers to explore relationships among video programming distributors.

## Appendix A: Additional Information

**Table A: Key Terms & Definitions**

Key Terms	Definition
Vertical Integration	Vertical Integration is the integration and merger of an upstream and a downstream firm in an industry with a supply chain. In this research paper, this refers to the merger of a programming distributor with a programming network.
Multichannel video programming distributors (MPVD)	MPVD is a service provider that provides video programming services. This includes cable television distributors, direct-broadcast satellite (DBS) providers, and wireline video providers.
Multiple system operator (MSO)	A multiple system operator is a subset of MPVD that specifically own local cable systems in the U.S. The two biggest MSOs currently are Comcast and Time Warner Cable.

**Table B: TMS Data Variable Explanation**

TMS Data Variable	Definition
hndid	Each headend's unique Identification number
chpos	All of the channels offered at each headend and the channel position or number within the offered bundle
chtype	Channel types are separated into basic bundle, extended basic bundle, and digital basic bundle
sumhhs	This is the number of households each headend can serve within the geographic location
msoid	For each cable headend, the data includes an unique MSO identification number that operates the respective cable headend

## Reference

- Time Warner Entertainment Co., L.P. v. FCC, 56 F.3d 151, 183 (DC Cir 1995).
- Satellite Broadcasting and Communications Association v. FCC, 275 F.3d 337 (4th Cir. 2001).
- Ammori, M. (January 2010). TV Competition Nowhere: How the Cable Industry Is Colluding to Kill Online TV. *freepress*, 1-40.
- Arango, T. (June 23, 2009). Cable TV's Big Worry: Taming the Web. *New York Times*.
- Chen, D. and D. Waterman. (2007). Vertical Ownership, Program Network Carriage, and Tier Positioning in Cable Television: An Empirical Study. *Review of Industrial Organization* 30(3), 227-251.
- Chipty, T. (2001). Vertical Integration, Market Foreclosure, and Consumer Welfare in the Cable Television Industry. *American Economics Review* 91(3), 428-453.
- Cooper, M. and Turner, D. (Sept. 29, 2007). The Negative Effect of Concentration and Vertical Integration on Diversity and Quality in Video Entertainment. *35th Research Conference on Communication, Information and Internet Policy (TPRC)*.
- Ford, G. and J. Jackson. (1997). Horizontal Concentration and Vertical Integration in the Cable Television Industry. *Review of Industrial Organization* 12, 501-518.
- Goolsbee, A. (2007). Vertical Integration and the Market for Broadcast and Cable Television Programming. *2007 FCC Media Ownership Studies*, Study 9 1-48.
- K. and R. Cavazos. (2005). Why is This Show so Dumb? Advertising revenue and Program Content of Network Television. *Review of Industrial Organization* 27, 17-34.
- Maddigan, R. and Zaima, J. (1985). The Profitability of Vertical Integration. *Managerial and Decision Economics*, 178-179.
- Ulhs, A. (2007, April 19). Digital Divide: the Issue of Net Neutrality. *Imprint Magazine*.
- Waterman, D. and A. Weiss. (1996). The Effects of Vertical Integration between Cable Television Systems and Pay Cable Networks. *Journal of Econometrics* 72, 357-395.

## Appendix B: Scripts

### Script 1: Merging Script (Source: Greg Crawford)

```

clear
set mem 1200m
set more off
capture log close

global root = "C:\Users\ayurukog\Documents\My Dropbox\Bloomberg\TMSData2010\ASCII files"
global root = "C:\Users\Robin Lee\Documents\My Dropbox\_myfiles\Cable Projects\Bloomberg\TMSData2010\ASCII files"
global root2007 = "C:\Users\Greg\Documents\My Dropbox\Cable TV Shared\Bloomberg.Data.Anal\TMS"
global root2007 = "C:\Users\ayurukog\Documents\My Dropbox\Cable TV Shared\Bloomberg.Data.Anal\TMS"
global root = "C:\Users\Greg\Documents\My Dropbox\Bloomberg\TMSData2010\ASCII files"
global tmp = "C:\Tmp"

log using "f02z.log", text replace

/*
* 2007 Data
*****
* Step 1: Merge together headend, headend-zip, and system (mso) info
u "$root2007\lineup_headend_prepped.dta", clear
keep if hndtype == 0 | hndtype == 2
su
codebook hndid
so hndid
sa "$tmp\aa.dta", replace

u "$root2007\lineup_hndzip.dta", clear
su
so hndid
sa "$tmp\bb.dta", replace

u "$tmp\aa.dta", clear
merge hndid using "$tmp\bb.dta"
tab _merge
keep if _merge == 3
drop _merge
codebook hndid
su
so sysid
sa "$tmp\cc.dta", replace

u "$root2007\lineup_system", clear
su
codebook msoid
so sysid
sa "$tmp\dd.dta", replace

u "$tmp\cc.dta", clear
merge sysid using "$tmp\dd.dta"
tab _merge
keep if _merge == 3
drop _merge
su

```

```

* Step 2: Define mso_ordered and Rename vars to match (as well as possible) 2010 data
foreach mso in cablevision charter comcast cox time_warner other none {
  gen `mso' = 0
}
replace cablevision = 1 if msoid == 4
replace charter = 1 if msoid == 592
replace comcast = 1 if msoid == 5
replace cox = 1 if msoid == 7
replace time_warner = 1 if msoid == 2
replace none = 1 if msoid == 549
gen othermso = 1 - cablevision - charter - comcast - cox - time_warner - none
gen mso_ordered = 0
replace mso_ordered = 1 if cablevision
replace mso_ordered = 2 if charter
replace mso_ordered = 3 if comcast
replace mso_ordered = 4 if cox
replace mso_ordered = 5 if time_warner
replace mso_ordered = 6 if othermso
replace mso_ordered = 7 if none
gen directv = (hndid == 16837)
gen dish = (hndid == 15164 | hndid == 18801)
replace mso_ordered = 8 if mso_ordered == 6
replace mso_ordered = 9 if mso_ordered == 7
replace mso_ordered = 6 if directv
replace mso_ordered = 7 if dish
label define mso_orderedlab2 1 "Cablevision" 2 "Charter" 3 "Comcast" 4 "Cox" 5 "Time Warner" 6 "DirecTV" 7 "Dish" 8 "Other
MSO" 9 "None"
label values mso_ordered mso_orderedlab2
tab mso_ordered
drop cablevision charter comcast cox time_warner none othermso directv dish

rename numzip syszip
replace hndname = reverse(substr(reverse(hndname),5,.))
rename hndname headend_location
rename hndid hndid_2007
rename msoid msoid_2007
*rename state state_2007
rename sysid sysid_2007
keep hndid_2007 hndzipcode msoname msoid_2007 headend_location mso_ordered sysid_2007

* Step 3: Change case on 2007 msoname in changecase_msoname_2007.xlsx + _updatebyhand
*tab msoname
*do "changecase_msoname_2007.do"
* Based on merge of above and comparison of owner names, make by-hand changes in the file below
do "changecase_msoname_2007_updatebyhand.do"

* Step 4: Bring in household zipcode data
su
so hndzipcode
sa "$tmp\ee.dta", replace

* Households: 106.8m in the raw data
u "$root\ziphhs", clear
drop if households == .
rename zipcode hndzipcode
so hndzipcode
sa "$tmp\ziphhs_sorted", replace

```

```

* Households: 105.3m in the merged data
u "$tmp\ee.dta", clear
merge hndzipcode using "$tmp\ziphhs_sorted"
tab _merge
keep if _merge == 3
drop _merge
sa "$tmp\ff.dta", replace

* Step 5: Get rid of satellite guys
*tab headend_location
drop if headend_location >= "DIRECTV Albany" & headend_location <= "DIRECTV West Palm Beach"
so hndzipcode msoname headend_location
rename mso_ordered mso_ordered_2007
sa "$tmp\gg.dta", replace

* Focus on the big channels
u "$root2007\lineup_lineup_prepped.dta", clear
drop if chtype == 5
rename hndid hndid_2007
keep hndid chpos chtype chref
so chref
sa tmp1, replace

u "$root2007\lineup_channel_new.dta", clear
keep chref chname
rename chname channel_name_2007
do "reconcile_channels_2007.do"
so chref
sa tmp2, replace

u tmp1, clear
merge chref using tmp2
tab _merge
keep if _merge == 3
drop _merge
tab kaganname_2007
drop if kaganname_2007 == ""
drop chref channel_name_2007
so hndid chtype
by hndid chtype: egen numchans = count(chpos)
sa "$tmp\ii_full.dta", replace
*duplicates drop hndid chtype, force
*drop chpos kaganname_2007
*sa "$tmp\ii.dta", replace

u "$tmp\gg.dta", clear
joinby hndid_2007 using "$tmp\ii_full.dta", unmatched(both)
tab _merge
keep if _merge == 3
drop _merge
so hndid chtype
*drop msoid_2007 state_2007
sa "$tmp\jj.dta", replace

so hndzipcode msoname headend_location
* Try to match on zipcode-headend_id = no luck!
*list hndzipcode state hndid msoname msoid, separator(0) string(15)
* Try to match on zipcode-headend_location

```

```

*list hndzipcode msoname headend_location , separator(0) string(24)
*/

/*
* 2010 Data
*****
u "$tmp\headend.dta", clear
su
codebook headend_id

* Step 2 (no step 1): Rename vars to match (as well as possible) 2010 data
foreach mso in cablevision charter comcast cox time_warner other none verizon att {
  gen `mso' = 0
}
replace cablevision = 1 if mso_code == 5190
replace charter = 1 if mso_code == 13890
replace comcast = 1 if mso_code == 8010
replace cox = 1 if mso_code == 7510
replace time_warner = 1 if mso_code == 8670
replace none = 1 if mso_code == 0
replace verizon = 1 if mso_code == 17762
replace att = 1 if mso_code == 17304
gen othermso = 1 - cablevision - charter - comcast - cox - time_warner - none - verizon - att
gen mso_ordered = 0
replace mso_ordered = 1 if cablevision
replace mso_ordered = 2 if charter
replace mso_ordered = 3 if comcast
replace mso_ordered = 4 if cox
replace mso_ordered = 5 if time_warner
replace mso_ordered = 6 if othermso
replace mso_ordered = 7 if none
gen directv = 0
gen dish = 0
*gen directv = (hndid == 16837)
*gen dish = (hndid == 15164 | hndid == 18801)
replace mso_ordered = 8 if mso_ordered == 6
replace mso_ordered = 9 if mso_ordered == 7
replace mso_ordered = 6 if directv
replace mso_ordered = 7 if dish
replace mso_ordered = 10 if verizon
replace mso_ordered = 11 if att
label define mso_orderedlab2 1 "Cablevision" 2 "Charter" 3 "Comcast" 4 "Cox" 5 "Time Warner" 6 "DirecTV" 7 "Dish" 8 "Other
MSO" 9 "None" 10 "Verizon" 11 "AT&T"
label values mso_ordered mso_orderedlab2
tab mso_ordered
drop cablevision charter comcast cox time_warner none othermso directv dish verizon att

rename zip_code hndzipcode
rename mso_name msoname
rename state state_2010
rename mso_code msoid_2010
rename headend_id hndid_2010
keep hndid_2010 hndzipcode msoid_2010 state_2010 msoname headend_location mso_ordered

* Step 4 (no step 3): Bring in household zipcode data
so hndzipcode
sa "$tmp\zz.dta", replace

```

```

* Households: 105.1m in the merged data
u "$tmp\zz.dta", clear
merge hndzipcode using "$tmp\ziphhs_sorted"
tab _merge
keep if _merge == 3
drop _merge
sa "$tmp\yy.dta", replace

* Step 5: Get rid of satellite guys
*so msoname
*by msoname: egen numhndzips = count(hndzipcode)
*gsort -numhndzips
*duplicates drop msoname, force
*list numhndzips numhndzips msoname msoid_2010, separator(0)
* DIRECTV = msoid_2010 == 17580
* Dish (Echostar) = msoid_2010 == 17680
drop if msoid_2010 == 17580 | msoid_2010 == 17680
so hndzipcode msoname headend_location
rename mso_ordered mso_ordered_2010
drop mso_ordered_2010 households msoid_2010 state_2010
sa "$tmp\xx.dta", replace

* Focus on the big channels
u "$tmp\lineups.dta", clear
drop if service_tier >= 3
destring channel_position, replace force
rename headend_id hndid_2010
rename station_num chref
rename service_tier chtype
rename channel_position chpos
keep hndid chpos chtype chref
so chref
sa tmp1, replace

u "$tmp\statrec.dta", clear
rename station_num chref
rename channel_name chname
keep chref chname
rename chname channel_name_2010
do "reconcile_channels_2010.do"
so chref
sa tmp2, replace

u tmp1, clear
merge chref using tmp2
tab _merge
keep if _merge == 3
drop _merge
tab kagannname_2010
drop if kagannname_2010 == ""
drop chref channel_name_2010
so hndid chtype
by hndid chtype: egen numchans = count(chpos)
sa "$tmp\ww_full.dta", replace
su

u "$tmp\xx.dta", clear
joinby hndid_2010 using "$tmp\ww_full.dta", unmatched(both)

```

```

tab_merge
keep if _merge == 3
drop _merge
so hndid chtype
sa "$tmp\vv.dta", replace

su
*/

*so hndzipcode msoname headend_location
* Try to match on zipcode-headend_id = no luck!
*list hndzipcode state headend_id msoname mso_code, separator(0) string(15)
* Try to match on zipcode-headend_location
*list hndzipcode msoname headend_location , separator(0) string(24)

* Merge Data
*****
/*
* Step 1: First merge only on hndzipcode msoname headend_location (jj2 & vv2, to save space)
* Call these goodobs

u "$tmp\jj.dta", clear
drop sysid_2007 msoid_2007
*hndid_2007
rename chtype chtype_2007
rename numchans numchans_2007
rename mso_ordered_2007 mso_ordered
so hndzipcode msoname headend_location chpos
sa "$tmp\jj1.dta", replace
keep hndzipcode msoname headend_location
duplicates drop hndzipcode msoname headend_location , force
sa "$tmp\jj2.dta", replace

u "$tmp\vv.dta", clear
drop hndid_2010
rename chtype chtype_2010
rename numchans numchans_2010
so hndzipcode msoname headend_location chpos
sa "$tmp\vv1.dta", replace
keep hndzipcode msoname headend_location
duplicates drop hndzipcode msoname headend_location , force
sa "$tmp\vv2.dta", replace

u "$tmp\jj2.dta"
merge hndzipcode msoname headend_location using "$tmp\vv2.dta"
tab_merge
keep if _merge == 3
drop _merge
gen goodobs = 1
so hndzipcode msoname headend_location
sa "$tmp\merged_thin.dta", replace

* Step 2: Having identified the goodobs, bring back in the rest of the info in each year

u "$tmp\merged_thin.dta", clear
merge hndzipcode msoname headend_location using "$tmp\jj1.dta"
tab_merge
* Losing about 58% of hndzip-msoname-headend-locations

```

```
keep if _merge == 3
drop _merge
so hndzipcode msoname headend_location chpos
* 0.3% of that 4-tuple are duplicates, drop those
duplicates drop hndzipcode msoname headend_location chpos, force
sa "$tmp\merged1.dta", replace
*/

u "$tmp\merged1.dta", clear
merge hndzipcode msoname headend_location chpos using "$tmp\vv1.dta"
tab _merge
drop _merge
so hndzipcode msoname headend_location
by hndzipcode msoname headend_location: egen mgoodobs = max(goodobs)
keep if mgoodobs == 1
drop goodobs mgoodobs
so hndzipcode msoname headend_location chpos
sa "$root\merged_all.dta", replace

log close
```

**Script 2: Network Carriage Percentage****Carriage Percentage all MSO 2007 2010.do**

```
clear
clear matrix
set mem 600m
capture log close
```

```
log using "Cablevision.top50carriage w allMSOs.supporting.log", text replace
```

```
*=====
```

```
* MERGE IN HOUSEHOLDS / HEADEND
```

```
*use "C:\Users\Jeffrey
Shen\Documents\Econ\Econ198\Research\Para136.Channel_changes\test_merged2007_2010whhs -
version 2.dta", clear
```

```
*Works
```

```
*use "C:\Users\Jeffrey
Shen\Documents\Econ\Econ198\Research\Para136.Channel_changes\merged2007_2010whhs.dta",
clear
```

```
*-----
```

```
*Final with 3 mill
```

```
*do "C:\Users\Jeffrey
Shen\Documents\Econ\Econ198\Research\Para136.Channel_changes\f04.comparison_of_lineups_hhs.s
upporting - channel_carriage full.do"
```

```
*-----
```

```
*-----
```

```
*With 21 mill
```

```
*use "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\2007_2010_MSO_Data\testing w fillers filled.dta", clear
```

```
*Sample Data
```

```
use "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\2007_2010_MSO_Data\2007_2010_MSO_merged_all.dta", clear
```

```
*Comcast Corporations has 1.95e+07 Households
```

```
*drop if msoname != "Comcast Corporation"
```

```
*Time Warner Cable has 1.00e+07 Households
```

```
*drop if msoname != "Time Warner Cable"
```

```
*Cablevision has 1958940 Households
```

```
replace msoname = "Cablevision Systems Corporation" if msoname == "Cablevision Services
Incorporated";
```

```
drop if msoname != "Cablevision Systems Corporation"
```

```
*Cox has 5245343 Households
```

```
*drop if msoname != "Cox Communications"
gsort +hndzipcode -hndid_2007 +chpos
```

```
replace hndid_2007 = hndid_2007[_n-1] if missing(hndid_2007)
rename hndid_2007 hndid
```

```
rename kaganname_2007 channel_name_2007
rename kaganname_2010 channel_name_2010
gen kaganname_2007 = ""
gen kaganname_2010 = ""
*drop kaganname_2007 kaganname_2010
do "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\Scripts\reconcile_channels_2007_top50channels.do"
do "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\Scripts\reconcile_channels_2010_top50channels.do"
```

```
merge m:m hndid using "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\2007_2010_MSO_Data\2007_MSO_all_sumhhs.dta"
```

```
drop if _merge != 3
*drop if msoname != "Comcast Corporation"
*drop if msoname != "Time Warner Cable"
drop if msoname != "Cablevision Systems Corporation"
*drop if msoname != "Cox Communications"
rename sumhhs hhs
*-----
```

```
foreach channel in AMC Bloomberg WEtv EEntertainment GolfChannel TravelChannel Hallmark CSPAN
AnimalPlanet Discovery DiscoveryHealth TLC TVGuide Bravo CNBC CNBCWorld Oxygen Syfy USA AandE
History MSNBC Weather FoxNews FSN FXNetwork SpeedChannel HGTV FoodNetwork ABCFamily
DisneyChannel DisneyXD LifetimeMovie Lifetime ESPN ESPN2 Cartoon CNN TBS TCM TNT TruTV
WGNAmerica BET CMT ComedyCentral MTV MTV2 Nickelodeon SpikeTV TVLand VH1{
```

```
display "`channel'"
*2007 Channels
gen `channel'_2007 = 1 if kaganname_2007 == "`channel'";
su `channel'_2007 if kaganname_2007 == "`channel'";
egen ch`channel'_2007 = sum(!missing(`channel'_2007)), by(hndid)
drop `channel'_2007
```

```
*2010 Channels
gen `channel'_2010 = 1 if kaganname_2010 == "`channel'";
su `channel'_2010 if kaganname_2010 == "`channel'";
egen ch`channel'_2010 = sum(!missing(`channel'_2010)), by(hndid)
drop `channel'_2010
}
```

```
duplicates drop hndid, force
```

```
*Sum of all Households
```

```
egen sumhhs = sum(hhs)
```

```
foreach channel in AMC Bloomberg WEtv EEntertainment GolfChannel TravelChannel Hallmark CSPAN  
AnimalPlanet Discovery DiscoveryHealth TLC TVGuide Bravo CNBC CNBCWorld Oxygen Syfy USA AandE  
History MSNBC Weather FoxNews FSN FXNetwork SpeedChannel HGTV FoodNetwork ABCFamily  
DisneyChannel DisneyXD LifetimeMovie Lifetime ESPN ESPN2 Cartoon CNN TBS TCM TNT TruTV  
WGNAmerica BET CMT ComedyCentral MTV MTV2 Nickelodeon SpikeTV TVLand VH1{
```

```
egen sum`channel'_2007 = sum(hhs) if ch`channel'_2007 >= 1, by(hndid)
```

```
egen sum`channel'_all_2007 = sum(sum`channel'_2007)
```

```
gen `channel'_percentage_2007 = (sum`channel'_all_2007*100)/(sumhhs)
```

```
egen sum`channel'_2010 = sum(hhs) if ch`channel'_2010 >= 1, by(hndid)
```

```
egen sum`channel'_all_2010 = sum(sum`channel'_2010)
```

```
gen `channel'_percentage_2010 = (sum`channel'_all_2010*100)/(sumhhs)
```

```
drop sum`channel'_2007 sum`channel'_all_2007 sum`channel'_2010 sum`channel'_all_2010
```

```
display "Carriage      `channel'      2007      "`channel'_percentage_2007
```

```
display "Carriage      `channel'      2010      "`channel'_percentage_2010
```

```
}
```

```
log close
```

**Script 3: Extract Network Carriage Percentage****Extract\_Household\_Carriage.pl**

```

#Jeffrey (Shih-kai) Shen
#Econ Honors Thesis

# Filepaths
$ownerfilepath = "C:\\Users\\Jeffrey
Shen\\Documents\\Econ\\Econ198\\Research\\Para136.Channel_changes\\readme Top 50 Channels w Owner
2007 and 2010 Merge Names.txt";

#Comcast
#$carriagefilepath = "C:\\Stata11\\Comcast.top50carriage w allMSOs.supporting.log";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Result
Data\\Comcast_Carriage_of_TimeWarner_top50networks_Result_2007_2010 w all MSOs Data.txt";

#Cox
#$carriagefilepath = "C:\\Stata11\\Cox.top50carriage w allMSOs.supporting.log";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Result
Data\\Cox_Carriage_of_TimeWarner_top50networks_Result_2007_2010 w all MSOs Data.txt";

#$outfilepath = "C:\\Users\\Jeffrey
Shen\\Documents\\Econ\\Econ198\\Research\\Para136.Channel_changes\\TimeWarner_Carriage_of_TimeWarne
r_top50networks_Result_2007_2010.txt";

#Cablevision
$carriagefilepath = "C:\\Stata11\\Cablevision.top50carriage w allMSOs.supporting.log";
$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Result
Data\\Cablevision_Carriage_of_TimeWarner_top50networks_Result_2007_2010 w all MSOs Data.txt";

#Time Warner
#$carriagefilepath = "C:\\Stata11\\TimeWarner.top50carriage w allMSOs.supporting.log";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Result
Data\\TimeWarner_Carriage_of_TimeWarner_top50networks_Result_2007_2010 w all MSOs Data.txt";

#Unknown
#$carriagefilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Probit\\Probit
Time Warner Fix\\TimeWarner.top50carriage w allMSOs.supporting.log";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Probit\\Probit Time
Warner Fix\\TimeWarner_Carriage_of_TimeWarner_top50networks_Result_2007_2010 w all MSOs Data.txt";

#Inputs
$MSO = "Time Warner"; #MSO interested #MSO = "Time Warner"
$MSO_carriage_carrier = "Cablevision"; #MSO System
$start_year = 2007; $end_year = 2010;

#Opening the Ownership File and Storing the Information in the Hash
open(ownerfile, "<", $ownerfilepath ) or die "Can't open ownership file\n";
%ownership = (); #Hash of ownership key: is the channel ownership: Company
while(<ownerfile>){
    chomp $_;

```

```

        if($_ !~ m/^(channel/)){

            @line = split(/\t/, $_);
            $key = $line[0]; #Key for the hash
            $ownership{"$key"} = $line[1];
#            print "$_\n";
            print "key: $key owner: $ownership{"$key"}\n";
        }
    }
close ownerfile;

#Printing for Outfile
open(outfile, ">", $outfilepath) or die "Can't open outfile\n";
print outfile "Carriage_Percentage"."\\t".$MSO."\\t"."Yr".$start_year."\\t"."Yr".$end_year."\\t";

foreach $ownership(sort keys %ownership){
    $key = $ownership;
    print outfile $key."\\t";
#    if($line[1] =~ m/$key/){
#        print "same\n";
#        $channel_list = $channel_list."1"."\\t";
#    }
#    else{
#        $channel_list = $channel_list."0"."\\t";
#    }
}
print outfile "\\n";

#Opening the Carriage File
print "\\nOpening the Carriage File\\n";

open(carriagefile, "<", $carriagefilepath) or die "Can't open carriagefile\n";
while(<carriagefile>){
    chomp $_;
    if($_ =~ m/^(Carriage/){
        print "$_\n";
        $row = rowFeature($_);
        print "Row: $row\n";
        print outfile "$row\n";
    }
}
close carriagefile;
close outfile;

sub rowFeature{
    print "This is in the sub: $_\n";
    @line = split(/\s+/, $_);

```

```

print "Array: $line[0] and $line[1] and $line[2] last $line[3]\n";

#Create 1 if MSO matches ex: CNN is owned by Time Warner so if $MSO = Time Warner, $MSO_match = 1
$MSO_match = ();
if($ownership{$line[1]} =~ m/$MSO/){
    $MSO_match = 1;
}
else{
    $MSO_match = 0;
}

#Create Year
$start_year_match = year_match($line[2], $start_year);

if($start_year_match == 1){
    $end_year_match = 0;
}
else{
    $end_year_match = 1;
}
print"2007: $start_year_match 2010: $end_year_match\n";

#Create channel_list
$channel_list = ();
foreach $ownership(sort keys %ownership){
    $key = $ownership;
    print "Ownership: $key: ";
    print "$ownership{$key}\n";

    if($line[1] =~ m/($key)$/){
        print "same\n";
        $channel_list = $channel_list."1"."t";
    }
    else{
        $channel_list = $channel_list."0"."t";
    }
}

$channel_matrix =
$line[3]."\t".$MSO_match."\t".$start_year_match."\t".$end_year_match."\t".$channel_list;
#    print "Channel Matrix: $channel_matrix\n";
return $channel_matrix;
}

sub year_match{
    if($_[0] =~ m/$_[1]/){
        return 1;
    }
    else{
        return 0;
    }
}
}

```

**Script 4: Generate Ownership****generate\_Ownership.do**

\*Generate Ownership Variables

clear

clear matrix

insheet using "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All MSOs\Result Data\Cablevision\_Carriage\_of\_TimeWarner\_top50networks\_Result\_2007\_2010 w all MSOs Data.txt"  
drop v56

\*Generate Ownership Variables

gen cablevision = 0

gen comcast = 0

gen cox = 0

aorder timewarner cablevision comcast cox

move carriage\_percentage cablevision

\*Generate Ownership Values

replace comcast = 1 if eentertainment == 1

replace comcast = 1 if golfchannel == 1

replace cablevision = 1 if amc == 1

replace cablevision = 1 if wetv == 1

\*Major Difference Fixing Cox

replace cox = 1 if animalplanet == 1

replace cox = 1 if discovery == 1

replace cox = 1 if discoveryhealth == 1

replace cox = 1 if tlc == 1

replace cox = 1 if travelchannel == 1

save "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All MSOs\Result Data\Cablevision carriage 2007 2010 All MSO Data.dta", replace  
regress \*

**Script 5: Create Probit Headends**

```
clear
clear matrix
set mem 600m
capture log close
```

```
log using "Cox.Probit.top50carriage w allMSOs.supporting.log", text replace
```

```
*=====
```

```
* MERGE IN HOUSEHOLDS / HEADEND
```

```
*Works
```

```
*use "C:\Users\Jeffrey
Shen\Documents\Econ\Econ198\Research\Para136.Channel_changes\merged2007_2010whhs.dta",
clear
```

```
*-----
```

```
*Final with 3 mill
```

```
*do "C:\Users\Jeffrey
Shen\Documents\Econ\Econ198\Research\Para136.Channel_changes\f04.comparison_of_lineups_hhs.s
upporting - channel_carriage full.do"
```

```
*-----
```

```
*-----
```

```
*With 21 mill
```

```
*use "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\2007_2010_MSO_Data\testing w fillers filled.dta", clear
```

```
*Sample Data
```

```
use "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\2007_2010_MSO_Data\2007_2010_MSO_merged_all.dta", clear
```

```
*drop if msoname != "Comcast Corporation"
```

```
*drop if msoname != "Time Warner Cable"
```

```
*replace msoname = "Cablevision Systems Corporation" if msoname == "Cablevision Services
Incorporated";
```

```
*drop if msoname != "Cablevision Systems Corporation"
```

```
drop if msoname != "Cox Communications"
```

```
gsort +hndzipcode -hndid_2007 +chpos
```

```
replace hndid_2007 = hndid_2007[_n-1] if missing(hndid_2007)
```

```
rename hndid_2007 hndid
```

```
rename kagannname_2007 channel_name_2007
```

```
rename kagannname_2010 channel_name_2010
```

```
gen kagannname_2007 = ""
```

```
gen kagannname_2010 = ""
```

```
*drop kagannname_2007 kagannname_2010
```

```
do "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\Scripts\reconcile_channels_2007_top50channels.do"
```

```
do "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\Scripts\reconcile_channels_2010_top50channels.do"
```

```
merge m:m hndid using "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All
MSOs\2007_2010_MSO_Data\2007_MSO_all_sumhhs.dta"
```

```
drop if _merge != 3
```

```
*drop if msoname != "Comcast Corporation"
```

```
*drop if msoname != "Time Warner Cable"
```

```
*drop if msoname != "Cablevision Systems Corporation"
```

```
drop if msoname != "Cox Communications"
```

```
rename sumhhs hhs
```

```
*-----
```

```
foreach channel in AMC Bloomberg WEtv EEntertainment GolfChannel TravelChannel Hallmark CSPAN
AnimalPlanet Discovery DiscoveryHealth TLC TVGuide Bravo CNBC CNBCWorld Oxygen Syfy USA AandE
History MSNBC Weather FoxNews FXNetwork SpeedChannel HGTV FoodNetwork ABCFamily
DisneyChannel DisneyXD LifetimeMovie Lifetime ESPN ESPN2 Cartoon CNN TBS TCM TNT TruTV
WGNAmerica BET CMT ComedyCentral MTV MTV2 Nickelodeon SpikeTV TVLand VH1{
display "`channel'"
```

```
*2007 Channels
```

```
gen `channel'_2007 = 1 if kaganname_2007 == "`channel'";
```

```
su `channel'_2007 if kaganname_2007 == "`channel'";
```

```
egen ch`channel'_2007 = sum(!missing(`channel'_2007)), by(hndid)
```

```
drop `channel'_2007
```

```
*2010 Channels
```

```
gen `channel'_2010 = 1 if kaganname_2010 == "`channel'";
```

```
su `channel'_2010 if kaganname_2010 == "`channel'";
```

```
egen ch`channel'_2010 = sum(!missing(`channel'_2010)), by(hndid)
```

```
drop `channel'_2010
```

```
}
```

```
duplicates drop hndid, force
```

```
*Sum of all Households
```

```
egen sumhhs = sum(hhs)
```

```
*foreach channel in Bloomberg CNN ESPN2 TBS TCM{
```

```
foreach channel in AMC Bloomberg WEtv EEntertainment GolfChannel TravelChannel Hallmark CSPAN
AnimalPlanet Discovery DiscoveryHealth TLC TVGuide Bravo CNBC CNBCWorld Oxygen Syfy USA AandE
History MSNBC Weather FoxNews FXNetwork SpeedChannel HGTV FoodNetwork ABCFamily
DisneyChannel DisneyXD LifetimeMovie Lifetime ESPN ESPN2 Cartoon CNN TBS TCM TNT TruTV
WGNAmerica BET CMT ComedyCentral MTV MTV2 Nickelodeon SpikeTV TVLand VH1{
```

```
egen sum`channel'_2007 = sum(hhs) if ch`channel'_2007 >= 1, by(hndid)
```

```
egen sum`channel'_all_2007 = sum(sum`channel'_2007)
```

```
gen `channel'_percentage_2007 = (sum`channel'_all_2007*100)/(sumhhs)
```

```
egen sum`channel'_2010 = sum(hhs) if ch`channel'_2010 >= 1, by(hndid)
```

```
egen sum`channel'_all_2010 = sum(sum`channel'_2010)
```

```
gen `channel'_percentage_2010 = (sum`channel'_all_2010*100)/(sumhhs)
```

```
*drop ch`channel'_2007 ch`channel'_2010
drop sum`channel'_2007 sum`channel'_all_2007 sum`channel'_2010 sum`channel'_all_2010

display "Carriage      `channel'      2007    "`channel'_percentage_2007
display "Carriage      `channel'      2010    " `channel'_percentage_2010
}

aorder *
keep ch* hhs sumhhs
drop channel_name_2007 channel_name_2010 chpos chtype_2007 chtype_2010

outsheet using "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All MSOs\Probit\Cox
Carriage headend data cut.txt", replace
keep hhs
save "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All MSOs\Probit\Cox hhs.dta"
outsheet using "C:\Users\Jeffrey Shen\Documents\Econ\Econ198\Research\All MSOs\Probit\Cox
hhs.txt", replace

log close
```

**Script 6: Probit Perl code****Extract\_Probit\_headend.pl**

```

#Jeffrey (Shih-kai) Shen
#Econ Honors Thesis

# Filepaths
$ownerfilepath = "C:\\Users\\Jeffrey
Shen\\Documents\\Econ\\Econ198\\Research\\Para136.Channel_changes\\readme Top 50 Channels w Owner
2007 and 2010 Merge Names.txt";
#Time Warner
#$headendfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\TimeWarner Carriage headend data cut.txt";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Result\\TimeWarner_Probit_Output.txt";
#
#$householdfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\TimeWarner hhs.txt";
#$householdoutfile = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\TimeWarner hhs_elongated.txt";

#Comcast
#$headendfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Comcast Carriage headend data cut.txt";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Result\\Comcast_Probit_Output.txt";
#
#$householdfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Comcast hhs.txt";
#$householdoutfile = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Comcast hhs_elongated.txt";

#Cablevision
#$headendfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Cablevision Carriage headend data cut.txt";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Result\\Cablevision_Probit_Output.txt";
#
#$householdfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Cablevision hhs.txt";
#$householdoutfile = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Cablevision hhs_elongated.txt";

#Cox
#$headendfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Probit\\Cox
Carriage headend data cut.txt";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\Result\\Cox_Probit_Output.txt";
#
#$householdfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Probit\\Cox
hhs.txt";

```

```
#$householdoutfile = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All MSOs\\Probit\\Cox
hhs_elongated.txt";
```

```
#Debugging
#$headendfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\TimeWarner Carriage headend data cut_Sample.txt";
#$outfilepath = "C:\\Users\\Jeffrey Shen\\Documents\\Econ\\Econ198\\Research\\All
MSOs\\Probit\\TimeWarner_Probit_Output_Sample.txt";
```

```
#Inputs
#$MSO = "Comcast"; #MSO interested
#$MSO = "Time Warner"
@MSO = ("Cablevision", "Comcast", "Cox", "Time Warner");
$MSO_size = @MSO;
$MSO_carriage_carrier = "Time Warner"; #MSO System
$start_year = 2007; $end_year = 2010;
```

```
#Opening the Ownership File and Storing the Information in the Hash
open(ownerfile, "<", $ownerfilepath) or die "Can't open ownership file\n";
%ownership = (); #Hash of ownership key: is the channel ownership: Company
while(<ownerfile>){
    chomp $_;
    if($_ !~ m/^(channel)/){

        @line = split(/\t/, $_);
        $key = $line[0]; #Key for the hash
        $ownership{"$key"} = $line[1];
    }
    print "$_ \n";
    print "key: $key owner: $ownership{$key} \n";
}
close ownerfile;
```

```
#Printing for Outfile
open(outfile, ">", $outfilepath) or die "Can't open outfile\n";
print outfile
"Carried." "\t". $MSO[0]. "\t". $MSO[1]. "\t". $MSO[2]. "\t". $MSO[3]. "\t". "Yr". $start_year. "\t". "Yr". $end_year. "\t";
```

```
foreach $ownership(sort keys %ownership){
    $key = $ownership;
    print outfile $key. "\t";
}
print outfile "\n";
```

```
open(headendfile, "<", $headendfilepath) or die "Can't open headendfile\n";
$offset = 2;
while(<headendfile>){
    chomp $_;
    if($_ =~ m/^ch/){
        print "$_ \n";
    }
}
```

```

@channel_line = split(/\t/, $_);
$channel_size = @channel_line;
#Taking out the ch and 2007 or 2010
@channel_line_fixed = @channel_line;
for($j=0; $j<$channel_size;$j++){
    $string = $channel_line[$j];
    $string =~ s/^ch//i;
    $string =~ s/_2007//i;
    $string =~ s/_2010//i;
    $channel_line_fixed[$j] = $string;
}
$num_channel = ($channel_size-$offset)/2;
print "Number of channels: $num_channel\n";
}
elseif($_ =~ m//^\d){
#    print "$_\n";
#    $row = Probit_rowFeature($_);
#    print "Row: $row\n";
#    print outfile "$row\n";
}
}
close headendfile;
close outfile;

sub Probit_rowFeature{
    print "This is in the sub: $_\n";
    @line = split(/\t/, $_);

    $size = @line;
    print "Start:";
    #For every entry in each headend row
    for ($i=0; $i<$size-$offset; $i++){
        #Carried
        print "\t$line[$i]";
        if($line[$i] >= 1){
            $Carried = 1;
            print outfile "\n$Carried";
        }
        else{
            $Carried = 0;
            print outfile "\n$Carried";
        }
    }
    #MSO Company
    for($j=0; $j<$MSO_size;$j++){
        print "Channel: $channel_line_fixed[$i]\n";
        print "Owner: $ownership{$channel_line_fixed[$i]}\n";
        print "MSO: $MSO[$j]\n";
        if($ownership{$channel_line_fixed[$i]} =~ m/$MSO[$j]/){
            $MSO_match = 1;
            print outfile "\t$MSO_match";
            print outfile "$channel_line_fixed[$i]";
        }
    }
}

```

```

        else{
            $MSO_match = 0;
            print outfile "\t$MSO_match";
        }
    }
    #Year
    if($channel_line[$i] =~ m/$start_year/){
        print outfile "\t1\t0";
        #
        print outfile "$start_year";
    }
    elsif($channel_line[$i] =~ m/$end_year/){
        #
        print outfile "\t0\t1";
        print outfile "$end_year";
    }
    }
    #Network
    for($k=0; $k<$channel_size-$offset;$k++){
        if($k == int($i/2)){
            #
            print outfile "\t1$channel_line_fixed[$k]";
            print outfile "\t1";
        }
        else{
            print outfile "\t0";
        }
    }
    }
    print"\n";

sub year_match{
    if($_[0] =~ m/$_[1]/){
        return 1;
    }
    else{
        return 0;
    }
}

#Make a sub that takes in array
open(householdfile, "<", $householdfilepath ) or die "Can't open householdfile file\n";
open(householdout, ">", $householdoutfile) or die "Can't open outfile\n";

while(<householdfile>){
    chomp $_;
    print "$_\n";
    if($_ !~ m/^(hhs)/){
        for($j=0; $j<$num_channel*2;$j++){
            print householdout "$_\n";
        }
    }
}

print "Number of channels: $num_channel\n";
print "Number of channels: $channel_size\n";

```