Vertical Integration in the U.S. Cable Industry*

Ayako Suzuki†

November 4, 2006

*I am truly indebted to Joseph Harrington and Matthew Shum. Additionally, Gregory K. Leonard, Robert Moffitt, Steffen Reinhold, Katie Winder, and Tiemen Woutersen provided valuable input to this study. I would also like to thank participants from the presentations at the following institutions: the Bank of Japan, the International Christian University, the International University of Japan, the Johns Hopkins University, and Osaka University.

†Institute of Social and Economic Research, Osaka University. Email: a-suzuki@iser.osaka-u.ac.jp. Tel: +81-6-6879-8581. Fax: +81-6-6878-2766. Address: 6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan.

Abstract

Theory shows that vertical integration has contrasting two effects, efficiency and foreclosure effects. This study empirically estimates the relative size of these two effects. Unlike previous studies, I focus on a single vertical merger in order to use a panel dataset, and estimate its average treatment effects on the several market outcomes. The findings suggest that there was a significant efficiency gain from the merger; the merged systems were found to carry affiliated networks more frequently; there was a larger price decrease in the merged markets. On the other hand, there was weak evidence of foreclosure. *JEL classification*: L10, L22, L40, C14.

1 Introduction

The causes and the consequences on market outcomes and consumer welfare of vertical integration have long been a controversial issue in economics and antitrust. Theory shows that vertical integration can promote efficiency by eliminating successive monopoly markups, internalizing service, and mitigating contractual problems between firms. It can also facilitate the strategic practice of market foreclosure, by which an integrated firm denies a rival access to downstream markets for the purpose of gaining market power. The first effect results in lower prices, higher sales, and higher consumer welfare, while the second effect raises the prices of both intermediate and final goods and thereby harms consumer welfare. Thus, assessing the effects of integration requires weighing the relative importance of these effects.

This study empirically examines the relative size of these effects towards understanding the welfare implications of vertical integration in the cable industry. This industry provides an ideal setting for such a study as it has experienced considerable backward integration between program service providers and cable system operators. The program service providers are upstream firms that provide cable networks to the cable system operators. The cable system operators are downstream firms that distribute these networks to consumers. The integration between these firms have raised concerns among the anti-trust authorities that consumers may be harmed if cable operators strategically foreclose competing programming networks. Such possibilities are of particular interest in the cable industry because program service providers do not usually have other means of distribution to consumers than selling to cable system operators, and each cable system faces little or no direct competition at the local level. As a result of such concerns, a restriction was proposed by the *1992 Cable Act* and implemented by the *Federal Communications Commission* (FCC). [1] In order to assess fairness of such a regulation, it is important to address the unanswered question of whether vertically integrated firms actually refuse to carry competing services on their distribution networks. However, this requires knowledge of the counterfactual outcome: What would have been the outcome if they had not been vertically integrated? This study examines this issue.

Using the system-level data from the *Television and Cable Factbook*, I study the effect of a vertical merger. The cable operators sell bundles of basic networks (often in the form of a basic service and extended basic services) and also offer the sale of individual premium networks tied to the basic bundle. Within this framework, cable operators choose prices, network offerings, and sizes of bundles to maximize profits . I estimate the effects of vertical integration on these endogenous variables.

Unlike previous studies, I focus on a single important vertical merger - *Turner Broadcasting* and *Time Warner* - and compare outcomes before and after the merger. This is opposed to previous studies that compare merged and non-merged cable systems at a point in time; here I examine the same systems before and after the merger. Using panel data enables me to estimate a dynamic response of the industry to the merger as opposed to cross sectional data by which one can only estimate an equilibrium outcome at a point in time. Furthermore, differencing outcomes across two time periods eliminates selection bias problems in terms of unobservables when they are time constant. Since participation in a merger is nonrandom, I adopt a non-experimental matching strategy to estimate the average treatment effect of the merger. Specifically, I employ the recently developed bias-adjusted matching method by Abadie and Imbens (2002).

The analysis shows that there is an efficiency gain from the merger of Turner Broadcasting and Time Warner. The merged systems carried affiliated networks more frequently than the non-merged systems; additionally, there was a larger price decrease in the merged markets, and the newly affiliated networks in the merged markets increased their numbers of subscribers to a greater than those in the non-merged markets. On the other hand, there is weak evidence of foreclosure.

The paper is structured as follows: the next section describes the structure of the US cable industry. Section 3 reviews the relevant theories and the previous empirical studies of vertical integration. Those familiar with the theory of vertical integration may want to skip the extensive discussion in Section 3.1. Section 4 describes the econometric method and data. The results are shown in Section 5 and Section 6 concludes the paper.

2 US Cable Industry

Both the division of the cable industry into a large number of independent local markets and the varying degree of vertical integration across these local markets make the industry an ideal setting to measure the effects of vertical integration on business practice. In this section, I review the structure and the degree of vertical integration in the industry.

In the industry, upstream firms are program service providers. They produce their own programming and/or acquire programming produced by others, and package and sell them as a network or group of networks to cable system operators for distribution to consumers. Under program service providers are downstream firms or cable system operators. They choose the number, the types, and the brand of networks to carry, set prices and sell to consumers in franchise areas. All operators sell a bundle of basic cable services for a single price per month, and most sell larger bundles of basic cable services, called extended basic services, and a la carte premium movie and sports services. Cable operators range in size from single-system cable operators with only a few dozen subscribers to *Multiple System Operators* (MSOs) that own many systems and serve millions of subscribers. Cable services are furnished in local franchise areas by one or two cable systems, and consumers cannot switch to another system that does not offer service within that area. Most franchise areas are in fact served by only one cable system and this deepens the concern for foreclosure.

During the last decade, cable operators have continued to consolidate and develop new ownership interests in program service providers. Hence, vertical integration observed in this industry is backward integration. More than 30 % of satellite-delivered national programming networks have been vertically integrated during the last decade.[2] Specifically, in 2003, 110 of the 330 existing satellite-delivered national programming networks were vertically integrated with at least one cable MSO. Vertical integration is not only associated with the largest cable system operators, but also the programming networks with the largest number of subscribers. In 2003, nine of the top 20 programming networks (ranked by subscribership) are vertically integrated with a cable MSO.[3]

The anti-trust concern is that vertically integrated cable operators may strategically favor their affiliated networks to the disadvantage of unaffiliated networks, either by the exclusion of rival networks from their program menu or, if the rivals are carried, by preferential pricing and promotion of their own networks. Such exclusion possibly increases sales (viewership) of the integrated networks within the operators' franchise areas and can ultimately benefit the integrated networks in unaffiliated franchise areas as rivals lose their competitiveness and possibly exit from the business. Hence, vertical integration may soften competition and harm consumers. On the other hand, vertically integrated firms may enjoy efficiency gains. Theory shows that vertical integration eliminates vertical externalities and benefits consumers. Furthermore, integration may reduce cost of negotiation between program services and cable operators, who negotiate and periodically renegotiate affiliate fees paid by the operator to the program service. Consequently, the effect of vertical integration in this industry is ambiguous. The purpose of this study is to clarify the relative size of these efficiency and strategic effects.

3 Theory and previous empirical studies

3.1 Theory

Theories of vertical integration are usually categorized by two approaches: neoclassical theories and organizational economics based theories (Joskow 2005). Neoclassical approaches explain vertical integration as a response to pre-existing market power problems or as a strategic move to create or enhance market power in upstream or downstream markets. The organizational economics based theories include the issues of transaction costs, incomplete contracts, and ownership that are crucial even in a world of perfect product-market competition.

As this study aims to understand the effect of vertical integration on consumer welfare, I reduce the theories into "efficiency" theories of vertical integration that are good for consumers, and anticompetitive "foreclosure" theories of vertical integration that are bad for consumers. Efficiency theories include both neoclassical and organizational economics-based theories that show vertical integration achieves efficiency gains, while foreclosure theory include those theories of neoclassical approaches that explain vertical integration as a strategic move to create market power.

3.1.1 Efficiency theories

Vertical externalities Neoclassical approaches explain vertical integration as a response to pre-existing market power problems. Pre-existing market power problems include various forms of vertical externalities. Three famous illustrations of vertical externality are given by Tirole (1988, Chapter 4) as describer below. Vertical integration can internalize these vertical externalities.

Double Marginalization: Suppose that the downstream firm's only decision is the final product price. The vertically integrated quantity q^m and consumer's price p^m are determined by $q^m = D(p^m)$, and p^m maximizes $(p - c)D(p)$, where $D(.)$ is the demand curve and c is the marginal cost for the upstream firm. Now consider the decentralized structure and the downstream firm's choice of the consumer price p under a linear wholesale tariff $T(q) = p_w q$. Assume that the upstream firm chooses the linear tariff first and the downstream firm chooses the consumer price second, and that the downstream firm is itself a monopolist in the retail market. The downstream firm maximizes its own profit, $(p - p_w)D(p)$. Because the downstream firm is a monopolist and because its marginal cost is equal to p_w, we have $p > p^m$ as long as the upstream firm charges above the marginal cost $(p_w > c)$, which is likely to be the case of the cable industry because the upstream market is an oligopolistic market. The final product price is higher in the decentralized structure than in the integrated one, because of these two successive mark-ups, and thereby downstream sales is lower in the decentralized structure. Vertical integration can internalize this externality, and result in higher sales, lower prices, and higher consumer welfare in the downstream market.

Downstream Moral Hazard: Downstream firms often provide services that make the manufacturer's good more attractive to consumers: trading stamps, free alterations, free delivery, credit, pre-sale information and so on. Cable operators also provide such services as free repair and replacement of equipment and a one-month trial period. We can call all these "promotional effort." To the extent that promotional effort affects the demand for the good, the upstream firm wants to encourage the downstream firm to supply it. The simplest way to do so would be to specify the level of promotional services in the contract. But such a contract would generally not be enforceable, as they cannot be measured precisely. Thus, incentives must be given to the downstream firm to overcome this moral-hazard problem.

Promotional services can be formalized as a real number, s. The consumer's demand is $q = D(p, s)$, where D decreases with p and increases with s. Assume that supplying a level s of services costs the downstream firm $\Phi(s)$ per unit of output (Φ increases with s), and that this cost can be observed only by the downstream firm. The vertically integrated consumer price (p^m) and services (s^m) maximize $[p - c - \Phi(s)]D(p, s)$. On the other hand, in the decentralized structure and for a linear price p_w, the profits are $(p_w - c)D(p, s)$ for the upstream firm and $[p - p_w - \Phi(s)]D(p, s)$ for the downstream firm. The upstream firm charges price $p_w > c$ in order to maximize its profit unless the upstream market is competitive. The downstream firm then maximizes its own profit with respect to p and s. There arises a distortion in consumer price as shown earlier. In addition, there is a distortion in services that resembles the distortion in consumer price. The downstream firm does not take into account the extra profit for the manufacturer associated with an increase in services, $(p_w - c)\frac{\partial D}{\partial s}$. Vertical integration can internalize this externality and increase the retailer's promotional effort. This will result in higher sales in the downstream market.

Input Substitution: Now suppose that the downstream firm is an industry that produces the final good from several inputs. This is particularly the case for the cable industry because the downstream firm gather several networks (inputs) and sell a bundle in this industry. To simplify, assume that the downstream firm uses two inputs: the upstream firm's good and a second intermediate good produced competitively at cost c'. Aside from the final price, the downstream firm must choose inputs x and x' to produce output $q = f(x, x')$. The two inputs are substitutes in the production function. Assume further that the technology exhibits constant returns to scale. The vertically integrated profit is $\Pi^m = \max_{x,x'}[P(f(x,x'))f(x,x') - cx - c'x']$, where $P(.) \equiv D^{-1}(.)$ is the inverse demand function.

Consider the decentralized structure in which the upstream firm of the first input x and the downstream industry have monopoly power.[4] Under linear pricing, the monopoly upstream supplier charges a wholesale price $p_w > c$, while the competitive upstream firm charge $p'_w = c'$ for the second input. Hence, the relative price of inputs for the downstream unit, $p_{w/}p'_w = p_w/c'$, exceeds their true relative price, c/c', for the vertical structure. The downstream unit thus substitutes toward the second input and consumes too little of the upstream monopolist's intermediate good. The downstream unit, when substituting toward the second input, does not take the upstream monopolist's marginal profit $(p_w - c)$ into account. Vertical integration internalize this externality. In the integrated case, the downstream monopolist pays a lower marginal price for the first input and therefore charges a lower price to consumers. As in the first theory, vertical integration results in a lower price and higher sales in the downstream markets. Furthermore, this shows that the downstream monopolist will increase the input purchased from the integrated upstream firm when they integrate.

Organizational economics based theories Organizational economics based theories, represented by Transaction Cost Economics (TCE) (Williamson 1975, 1983, 1985) and property right theories (Grossman and Hart 1986, Hart and Moore 1990), also show that vertical integration has efficiency improving effects. The crucial aspect of these theories is that even though the supplier and the buyer may select each other ex ante in a pool of competitive suppliers and buyers, they end up forming an ex post bilateral monopoly in that they have an incentive to trade between each other rather than with outside parties. [5] The theory also emphasizes the imperfection of contracts. When contracts are imperfect, parties often find themselves in the situation where bargaining occurs but may create inefficiency.

Tirole (1988, Chapter 2) raises an example in which bargaining creates ex post inefficiency. Assume that there are two periods: $t = 1$ (ex ante) and $t = 2$ (ex post). At the beginning of period 2, the two parties learn how much they will gain from trading in period 2. If they choose to do so, they trade one unit of an indivisible good. Thus, the volume of trade is either 0 or 1. The value of the good to the downstream firm is v, and its production cost to the upstream firm is c. Hence the gains from trade (if any) to be split between the two parties are equal to $v - \dot{c}$. If p is the trading price, the downstream firm's surplus from the relationship is $v - p$ and the upstream firm's surplus is $p - c$. In the absence of trade, the surplus is zero for the two parties. Further assume that some bargaining occurs in period 2 to determine whether to trade and at what price. In this kind of situation, it is well known that asymmetric information yields bargaining inefficiencies. Suppose that v is known to the downstream firm only, and that the upstream firm's beliefs about v are represented by a cumulative probability distribution $F(v)$ with density $f(v) > 0$ on an interval $[\underline{v}, \bar{v}]$. If the upstream firm offers p, the downstream firm accepts if and only if $v \geq p$. Thus, the probability of trading is $1 - F(p)$ and the upstream firm's expected profit is $(p-c)[1-F(p)]$. Maximizing this profit with respect to p yields an optimal price p^* where $p^* > c$. On the other hand, the efficient trade level arises when $p = c$; thus bargaining creates too little trade. Vertical integration should eliminate this inefficiency by internalizing the bargaining process, and increase the trade between the vertical parties.

This is just one example and any locked-in situation in which one party can act opportunistically could lead to inefficiency. The vertical relationship in the cable industry is likely to create such a locked-in situation because once a network is offered to consumers, it would be difficult or costly for cable operators to withdraw it from their services (i.e. the switching cost for cable operators may be high). Hence, upstream firms may behave opportunistically when they must re-negotiate the contract and downstream firms may be discouraged from carrying upstream networks.

Property right theories argue that the relevant comparison is between a contract that allocates residual rights to one party and another while TCE approach compares the nonintegrated and integrated outcomes. That is, integration shifts the incentives for opportunistic behavior from one party to the other, but it does not remove these incentives. According to these theories, backward integration should be observed in the cable industry should be observed when the specific investment of downstream firms is more important than that of upstream firms. Backward integration will increase the downstream firms' specific investment up to the second best level.

The above "efficiency" theories show that vertical integration is welfare increasing. The neoclassical theory shows that vertical integration eliminates vertical externalities arising from non-competitive markets and results in lower consumer prices and higher sales. Furthermore, as the input substitution example shows, if vertical integration can realize such efficiency gains, then the integrated downstream firms will substitute more towards the input from their affiliated upstream firms. Organizational economics based theories show that vertical integration reduces transaction costs, which involve the direct costs of writing, monitoring and enforcing contingent contracts as well as the costs associated with the ex ante investment and ex post performance inefficiencies, and results in higher ex-ante investment and higher trade of input. Hence, in this study, we expect to see that in the integrated franchise markets, prices for subscribers are lower, penetration rates are higher and the affiliated networks are carried more. However, we may not observe such differences between integrated and non-integrated markets if the industry has already achieved efficiency by other contractual means. It is well-known that non-linear pricing for inputs can eliminate vertical externalities (Tirole 1988 Chapter 4). For example, a two-part tariff can eliminate double mark-ups and a quantity-forcing contract is a sufficient instrument to achieve the vertically integrated profit when there is a downstream moral hazard problem.

3.1.2 Foreclosure theories

Market foreclosure are commercial practices that reduce the buyers' access to a supplier or limit a supplier's access to a buyer. In this study, I examine the latter issue, or downstream foreclosure. The concern here is that vertically integrated cable operators may refuse to carry unaffiliated cable networks in their franchise markets.[6] Here, I describe two classes of theories that show foreclosure can hurt consumers.[7] The first example shows that foreclosure increases the monopolization of one of the two sectors in the vertical relationship. The second one shows that foreclosure is used to form a barrier to entry in upstream markets.

Foreclosure for monopolization Hart and Tirole (1990) has a simple model called *Scarce Needs* that explains how foreclosure leads to monopolization. The model consists of two potential suppliers or upstream firms, U_1 and U_2, and two potential buyers or downstream firms, D_1 and D_2. The downstream firms compete on the product market and sell perfect substitutes. The upstream firms produce the same intermediate good at the same constant marginal cost. In such an situation, an upstream firm may merge with a downstream firm to ensure that the downstream firm purchases its supplies from this upstream firm rather than from others. In their model, non-linear pricing is assumed to be available and there is not excessive mark-ups and they suppose that a constant fraction $\beta < 1/2$ of the surplus goes to upstream firm. Hence, if neither of the firms integrate (i.e. each downstream firm trades with both upstream firms), each upstream firm gains profit of $U^{NI} = 2\beta\pi^d$, and each downstream firm gains profit of $D^{NI} = (1 - 2\beta)\pi^d$, where π^d is the Cournot level profit in the downstream market. Let q^* be the Cournot level output. Then the total output in the downstream market in this non-integrated case is $2q^*$.

Now suppose U_1 and D_1 merge, and D_1 buys all its input from U_1. Thus, U_1 gains a valuable trading opportunity and U_2 loses one (Scarce Needs refers to the fact that D_1 and D_2 have limited input requirements). If U_2 remains in the industry, the only effect of the merger is to increase the $U_1 - D_1$ share of industry profit and reduce the $U_2 - D_2$ share. Given that U_2 is as efficient as U_1, there is no reason for U_1 to restrict its supply to D_2. Hence, the profit for integrated firm $U_2 - D_1$ becomes $V^{PI} = (1 + \beta)\pi^d - E$, while the profit for non-integrated upstream firm U_2 is $U^{PI} = \beta\pi^d$ and the profit for D_2 is $D^{PI} = (1 - 2\beta)\pi^d$, where E is the cost for integration. The total output is still $2q^*$ and consumers are not affected at this point. However, if the reduction in the profits of U_2 causes it to quit the industry,[8] U_1 is left as the only supplier and given that it is merged with D_1, it will be able to use this power to completely monopolize the market ex post (as part of a merged firm, it has no incentive to supply D_2). Thus, total quantity supplied will fall to the monopoly level and the price consumers pay will rise.

Although a cable system is usually a monopoly in its franchise market, there are other video programming distributors such as direct broadcast satellite (DBS) services and hence, this theory with an oligopolistic downstream market can be applied to the cable industry.[9]

If vertically integrated firms in the cable industry actually exercise foreclosure, make rival networks exit from their markets, and monopolize the market for the delivery of video programming that should include satellite services and other distributors, integration harm consumers by raising the prices.

Foreclosure as barriers to entry It has been shown that exclusivity contracts such as exclusive dealing form a barrier to entry. Such contracts force new upstream firms to set up their own distribution networks. Thus, new upstream firms are less inclined to enter. Because new entry would have induced competition and such competition would have benefitted consumers, such an exclusive dealing cannot be favored by consumers. Downstream foreclosure due to vertical integration will act similarly (Aghion and Bolton 1987). This theory predicts that there are less entrants in vertically integrated upstream markets in addition to higher prices in the integrated markets.

3.2 Empirical literature

There is little research that has empirically examined the relative importance of these various effects of vertical integration,[10] and due to data availability, studies have concentrated on the cable industry (Chipty (2001), Waterman and Weiss (1996)) and the gasoline industry (Vita (2000), Gilbert and Hastings (2001)). The results from these studies are consistent with either a foreclosure or an efficiency theory of vertical integration. For example, in the case of the cable industry, Chipty (2001) finds that vertically integrated cable distributors are more likely to exclude rival cable programming networks and favor their affiliated networks. Overall, however, consumers are not harmed by this behavior since the resulting changes in prices and product variety appear either not to harm or to benefit consumers overall.

Waterman and Weiss (1996) examine the same issues for premium cable networks and find extensive evidence of exclusionary behavior by vertically integrated firms. They find little evidence of any downstream price effects but do find that sales (penetration) are higher for affiliated programming services carried by vertically integrated firms than for non-integrated firms. This suggests that vertical integration leads to increased downstream sales effort associated with distributor-owned programming services.

The departure of my study from previous studies lies in econometric methods and data. Specifically, I employ a non-experimental matching strategy as opposed to OLS or heavily parameterized binary choice models (such as Probit or Logit), which have been used in previous studies. The method used here may be preferred to those regression estimators when there are differences in the covariate distributions for treated and control groups, and I will show this is likely to be the case here because the mean values of covariates differ much across the groups. Furthermore, I focus on a single vertical merger and use panel data that consists of the years before and after the merger. This is opposed to the previous studies that use cross sectional data. Differencing variables across these two time periods eliminates selection bias problem in terms of unobservables when they are time constant. Using panel data also enables me to estimate a dynamic response of the industry to the merger as opposed to cross sectional data by which one can only estimate an equilibrium outcome at a point in time.

The next section describes the econometric method and the data.

4 Econometric Method and Data

Recently there has been much interest and studies in econometric method for estimating the effects of active labor market programs. The conventional impact of interest is the average impact of the treatment, or the indicator of whether an individual participated such an program or not. With experimental data, this average effect of the treatment can be obtained by comparing the sample averages of the treated and the control groups. Such an experimental data may be, however, difficult to obtain in reality. Hence, the recent econometric literature has explored the strategies for non-experimental data and developed methods to estimate the average treatment effect under the unconfoundedness assumption - "the assumption that the receipt of treatment is independent of the potential outcomes with and without treatment if certain observable covariates are held constant" (Imbens 2004).

Applications of such econometric methods should not be limited to labor markets. In this study, I estimate the average effects of the merger, or a binary treatment of whether a cable system is vertically integrated or not, on downstream market outcomes. Because the participation for this merger is not random, I adopt such a non-experimental strategy. Specifically, I use the method of matching, which explicitly selects a group of individuals that are "comparable" to the treatment group of cable systems. In this method, two groups are nonparametrically matched on common values of X_i, or observed characteristics of a market i and the outcomes of the matched-comparison group are used to construct the counter-factual outcomes for the treated individuals, corresponding to their outcome had they not received the treatment. I choose the bias-adjusted matching estimator developed by Abadie and Imbens (2002).

The matching methods, like other non-experimental strategies, assume unconfoundedness. This assumption does not hold when there are unobservable differences between the treated and control groups. The factors such as a cable system's managerial and sales skills seem to violate the assumption. To mitigate this problem, I incorporate the difference-in-differences method into the matching method. In order to have pre and post treatment data, I focus on a single vertical merger, the one between Turner Broadcasting and Time Warner in 1996 and collect data for 1995 and 2000 so as to use the difference-in-differences method. If the unobserved factors are time invariant, the difference-in-differences matching estimators can eliminate biases due to such factors. Furthermore, I conduct an exact match on state to capture a state-level transitory shock.

The following subsection explains the method and the data.

4.1 Method and implementation

4.1.1 Treatment

Definition The definition of treatment depends on the question of interest. In this study, I use the 1996 merger between Turner Broadcasting and Time Warner. I define my treatment as a single vertical merger in order to use the panel data that consists of before (1995) and after (2000) the merger.

As a result of this merger, all the cable systems that were owned by Time Warner in 1995 are now vertically integrated with Turner Broadcasting. In addition, there are some cable systems that were purchased by Time Warner after the merger occurred. That is, those systems were not owned by Time Warner in 1995, but they are owned by Time Warner in 2000 and hence vertically integrated with Turner Broadcasting. I exclude these cable systems from the treated group and limit the group for the systems that were already owned by Time Warner in 1995. The reason I exclude the systems owned by Time Warner after 1996 is because the pre-and-post difference of market outcomes for these cable systems may reflect more than the effect of vertical merger. They may have changed behavior not because they were integrated with Turner Broadcasting, but because they are now owned by Time Warner and gained horizontal market power. In order to assess the pure effect of vertical integration, I define my treatment in this study such that the treatment status is 1 if a system is vertically integrated with Turner Broadcasting without changing ownership during the examination period.

Turner Broadcasting and Time Warner This merger was between one of the biggest cable programming firms and one of the biggest cable operators. Turner Broadcasting is a major producer of news and entertainment product around the world and a provider of programming for the basic cable industry. It is home to an array of cable TV networks, including multiple CNN news properties, familiar entertainment networks such as TBS superstation (TBS), Turner network TV (TNT), Cartoon Network (TOON), and Turner Classic Movies (TCM), and specialized networks such as Turner South and Boomerang. After the merger, these networks were owned by the new merged company. For convenience, I call this new merged firm Turner-Time-Warner. The merger value was $7.5 billion.

The merger at first faced the opposition of the Federal Trade Commission (FTC) because of concern over anti-competitiveness in the cable industry. Tele-Communications Inc. (TCI), the biggest cable operator, owned 21% of Turner and would have owned about 9% of the merged company. Regulators were concerned that the direct link would reduce TCI and Time Warner's incentives to compete against each other. The merger was approved on July 17, 1996 after the companies agreed to amendments that limit the role of major shareholder

TCI in the merged company and a pledge that Time Warner will act in a non-discriminatory way toward rival cable operators and programmers.[11]

The advantage of using this case for our study is that it caused ownership movements of many networks that have different characteristics. Networks such as TNT, TBS and CNN Headline News, sold in this deal, are well-established popular networks (established in 1983, 1976 and 1982 respectively), while Cartoon network (TOON), Turner Classic Movies (TCM) were relatively new networks (established in 1992 and 1994). Such variation will prove to be instrumental in our analysis. Furthermore, the genre of CNN, TOON, and TCM are clearly defined (news, children network and movies respectively) and it makes it feasible to measure the degree of competition these networks are facing.

4.1.2 Data

For this study, I use two main data sources: *Television and Cable Factbook* and *Census*. Television and Cable Factbook compiles survey responses from all existing cable franchises. For each of the cable franchises in the United States, the data contains information on the system's owner, date franchise was awarded, total miles of cable plant, channel capacity and the number of homes with access to cable (referred to as homes passed) within the system's franchise area. The data also contain a description of the system's program offerings, monthly price and the number of subscribers for each cable service. I use volumes 1996 and 2001 of this Factbook which provide data for 1995 and 2000 respectively. [12] In this study, I focus on the top eight states in terms of the percentage of the merged systems among all the systems in a state: NY, NC, NH, OH, MS, WI, FL, MA[13]. I then have 1391 cable systems to examine in which the number of observations in the treated group (i.e. the number of the merged cable systems) is 348 for each year.

Dependent Variables Table 1 shows the summary statistics of the dependent variables. $Treated$ and $Control$ are the mean values in the treated and controlled groups respectively. DID refers to the difference-in-differences treatment effect on the treated without controlling the covariates.

The first four variables are the outcomes in the downstream markets. They are, respectively, changes in the total number of Turner Broadcasting's networks, in the number of non Turner Broadcasting's networks, in monthly subscription fee per network, and in the penetration rate (the number of subscribers divided by the number of homes passed), all for the basic service and extended basic services. It can be seen that, on average, per-network prices decreased, penetration rates increased, and the number of Turner Broadcasting's networks increased in both merged (treated) and non-merged (controlled) markets. The degree of such change differs between merged and non-merged markets. In merged markets, per-network price decreased less while penetration rate, the number of networks and the number of Turner Broadcasting's networks increased more compared with non-merged markets.

The next five variables are the Turner Broadcasting's networks. These variables are differences of the indicator variables in 2000 and 1995. An indicator variable for network N in market m at time t takes the value 1 if network N is carried in market m at time t. Hence, the mean value of an indicator variable at time t expresses the national market share of that network. Therefore, the average of difference in indicator variables in two time points indicates the change in national market share of the network. TBS, $Cartoon$, CNN, TCM and TNT refer to *TBS superstation*, *The Cartoon Network*, *CNN*, *Turner Classic Movies*, and *Turner Network TV* respectively. It can be seen that every Turner Broadcasting's network increased its national market share during this period.

The next two variables are networks owned by Time Warner. They are not Turner Broadcasting's networks, but partially owned by Time Warner. In that sense, these networks are affiliated to the new merged firm, Turner-Time Warner. *Comedy* and *CourtTV* refer to *Comedy Central* and *Courtroom Television Network* respectively. Comedy Central provides original comedy shows, often in a live format. The network is jointly owned by Time Warner and Viacom, in a joint venture known as Comedy Partners. Court TV's daytime programming is devoted to broadcasts of America's legal proceedings while its evening programming block features talk shows, documentary features and drama series. Again, both networks increased their national market shares.

The rest of the variables are the rival networks of the Turner Broadcasting's networks. These are the networks that have the same feature of the Turner Broadcasting's networks. The first three networks refer to *CNBC*, *Fox News Channel* and *MSNBC* respectively and are news channels. I chose these networks as rival networks of CNN. The networks *Disney Channel* and *Nickelodeon* are children's networks. They are chosen as rival networks to The Cartoon Network. The next three networks refer to *American Movie Classics*, *Independent Film Channel* and *SCI FI Channel* respectively. They are movie channels and rivals of TMC. The last three networks refer to *Fox Family Channel*, *FX* and O*dyssey Network* and are picked as rival networks to TNT and TBS superstation. These networks provide a diverse slate of family entertainment programs such as movies, sports, comedies and children's programmes, and I categorize them as an entertainment network. These networks increased their national market share as well.

The DID estimator shows that the merger affected positively for most of the variables.

Our purpose is to see if this is also the case after we control the covariates.

Covariates The matching variables can be partitioned into two sets: $X = (V, Z)$ where V are variables in the outcome equations and Z are variables in the participation (merger) equation; Z variables determine the treatment. To identify the variables V that affect the downstream outcomes in the cable industry, I modify the structural model of Waterman and Weiss (1996) in Appendix B While Waterman and Weiss focused on pay services of the cable systems, I focus on basic services because of the feature of Turner Broadcasting. The model for basic services may be more complicated because of the fact that they are bundles of many networks. As can be seen in the Appendix, the model induces the equations of the downstream outcomes that depend on ownership characteristics that affect the input cost, non-input cost, capacity to carry networks, demographic demand factors and some unobservable variables. We use the owner's horizontal size as ownership characteristics. Non-input cost depends on the size and efficiency of systems. We proxy efficiency with age of a system and population density of local markets. The unobservable factors should be differenced away if they are time-invariant. Table 2 summarizes the definitions of these variables. Most of them overlap with the variables used in Chipty (2001).

The variables that affect the participation of the treatment should be carefully considered. As stated before, although a system does not make a decision to vertically integrate, whether or not to sell itself to Time Warner is a system's decision. We assume that such a decision depends on the system's profitability, and hence depends on the endogenous variables in the downstream market (see equation (7) in the model in Appendix B). Therefore, we again use the variables that determine the downstream outcomes described above for the variables in the participation equation or Z. Because these variables in the participation equation should be determined pre-treatment, the 1995 data are used for Z variables. On the other hand, the difference between 1995 and 2000 are used for V variables.

To implement the matching procedure, variables that depend on whether or not treatment is received cannot be used as the matching variables. The Z variables can clear this criteria because they are variables of pre-merger. However, some differenced variables (i.e. V) should be used with caution. Specifically, system and owner characteristics such as changes in number of homes covered, channel capacity and miles of plant may depend on whether the system is involved with merger as the merger may enable systems to invest to improve these variables. Hence, I will exclude these variables from the matching variables. The demographic variables should be exogenous to the merger status and can be used for matching variables. Table 3 summarizes the matching variables.

Table 4 shows the summary statistics of the matching variables described above. The last two columns show the mean values of the treated group and the controlled groups respectively. As can be seen, the two groups have different characteristics pre treatment. For example, the mean values of owner size, density, home passed, and length of plant are much different in the two groups. Assuming that these variables proxy size and efficiency of the systems (see Table 2), the treated group consists of the systems that are larger and efficient. On the other hand, the demographic characteristic of the percentage of young viewer is not so different between the two groups.

4.1.3 Estimator

What we are interested in is the average effect of this vertically integration on market outcomes in the integrated markets. Because each observation can receive one and only one treatment in each time period, one of the outcomes in the two states is not observed. Missing data is a fundamental problem in any impact evaluation study. With experimental data, this average effect of the treatment can be obtained by comparing the sample averages of the treated and the control groups. Upon first glance, the treatment here, the indicator of whether a system is vertically integrated or not, seems to have the feature of experimental data: cable systems happen to find themselves to be vertically integrated because the decision for vertical integration was made by cable operators, and not by individual cable systems.[14] However, while the decision of vertical integration is exogenous to individual local cable systems, the decision of whether to be owned by the large multiple system operators (MSO) is not. Whether to sell its company to the MSOs is a decision of a cable system, and whether to purchase a cable system is a decision of cable operators. Since such decisions may depend on characteristics of the systems, an endogeneity problem arises.[15] In this reason, I adopt a non-experimental strategy to estimate the average treatment effect of the merger.

The simplest non-experimental strategy may be to use the linear regression estimator. Such regression estimators impute the missing potential outcomes using the estimated regression function. However, the simple regression estimators may be very sensitive to differences in the covariate distributions for treated and control groups because these estimators rely heavily on extrapolation (see Imbens (2004), pp 13). As seen in Table 4, our covariates have significantly different mean values across the groups. Hence, in this case, the linear estimator may suffer from these biases.

Matching estimators also impute the missing potential outcomes, but using only the outcomes of the opposite treatment group that have the closest characteristics. In this method, two groups are nonparametrically matched on common values of X_i, or observed characteristics of a market i and the outcomes of the matched-comparison group are used to construct the counter-factual outcomes for the treated individuals, corresponding to their outcome had they not received the treatment. In this sense, the estimators rely less on extrapolation, and biases from it should be smaller. Furthermore, our data set has many observations in the controlled group and it can be raised as one of the reasons that favor marching methods here. [16]

The matching estimators, however, generate biased results if the matches are not exact. Abadie and Imbens (2002) shows that the bias of the matching estimator may dominate the variance if the dimension of the covariates is large and suggests the additional bias correction.[17] Since I have many continuous covariates, such bias may be serious in this study.[18] Therefore, I follow their bias adjusted matching estimator that adjusts the means of outcomes in the set of matched observations by differences in the covariates between that set and the overall comparison group (see Appendix $A.2.2$). This estimator incorporates both of the consistency property of the matching and the desirable variance property of regression.

The bias adjusted matching estimator, like other non-experimental strategies, assumes unconfoundedness: the assumption that the receipt of treatment is independent of the potential outcomes with and without treatment conditional on observable covariates. This assumption does not hold when there are unobservable differences between the treated and control groups that affect the outcomes. The factors such as a cable system's managerial and sales skills seem to violate the assumption. To mitigate this problem, I incorporate the difference-in-differences method into the matching method (see Appendix $A.2.3$ for the final estimator I use, the bias-adjusted matching estimator with difference-in-differences). If the unobserved factors are time invariant, the difference-in-differences matching estimator can eliminate biases due to such factors. Only the bias that is caused by the differences in time variant unobservables between the two groups cannot be wiped out by this method. It seems to be extremely difficult to find the appropriate instrument variables in this case because if the participation of the merger depends on each cable system's profitability, any variable that may affect the merger decision should also affect the outcome variables. To mitigate the biases due to time variant unobservables, I conduct an exact match on state. This should capture a transitory shock if it is state level.

5 Empirical Results

Table 5, 6 shows the estimated average impact of the merger between Turner Broadcasting and Time Warner in the Time Warner's downstream markets. Here, the treatment is defined as "*being vertically integrated with Turner Broadcasting without changing its ownership by Time Warner*". The average treatment effects on the treated are estimated by the bias-adjusted matching method with difference-in-differences. For matching, it uses the Euclidean norm to measure the distance between different values for he covariates, after normalizing the covariates to have zero mean and unit variance. The units are exactly matched on the state. For the bias adjustment, the regression uses all the matching variables. The standard errors are estimated using the variance estimator in Abadie and Imbens (2002)[19].

The problem of matching methods remains that there is not a theoretical justification for determining an optimal number of matches. Abadie and Imbens (2002) shows that a simple matching method produces radically different estimates when the number of matches changes. If we increase the number of matches the quality of the matches goes down, while the efficiency loss decreases.[20] Hence in general, there is a trade off between efficiency and biasdness when we increase the number of matches, although Abadie and Imbens shows that the bias-adjusted matching estimator shows much more robust behavior when the number of matches changes. In their Monte Carlo simulation, Abadie and Imbens shows that the bias-adjusted matching estimator best performs when the number of matches around 4 to 16, in terms of the root-mean-squared-error and the median-absolute-error.[21] Hence, I show the results with 4 to 10 matches here. The results with different number of matches are also available.

5.1 Efficiency effect

The first category of Table 5, 6 shows the treatment effects on the outcomes of the downstream market: the changes in the number of Turner Broadcasting's networks in the basic/extended basic cable services, in the number of non-Turner Broadcasting's networks, in the price per network and in the penetration rate. The second category of Table 5, 6 shows the average treatment effects on the affiliated networks. The first five networks are Turner Broadcasting's networks and thus were owned by the merged cable operator (Time Warner) as a result of the merger. The last two networks, *Comedy Central* and *Court TV*, are not Turner's networks, but were owned by Time Warner before the merger.

Except the news channel, *CNN,* that received special restrictions by the FTC, the signs of the average treatment effects are all as predicted by the efficiency theories. That is, the treatment effects on the affiliated networks and that on the number of Turner's networks are positive (except TNT with M=4) implying that the merged firm substitutes more towards the inputs of the affiliated upstream firms. Similarly, as consistent with the efficiency theory, the treatment effect on the consumer price is negative, that on the penetration rate (as sales) is positive though the latter is not significant.

The positive average treatment effects on the affiliated networks are significant for TBS superstation and Turner Classic Movies at least at 5% level. Specifically, the result shows that TBS Superstation increased its market share 11.3 to 12.7 % more in the merged markets than in the non-merged markets. The relatively new network, Turner Classic Movies, also benefits from the merger. Again, the market shares of this network increased about 9.3 to 14.5% more. Reflecting this, the average treatment effect on the number of Turner's networks is positive and significant. The merged systems increased the number of Turner Broadcasting's networks about 0.29 to 0.44 networks more after the merger. This estimated positive and significant treatment effects on the affiliated networks are consistent with any efficiency theory. That is, these results suggest that the merger either reduces vertical externalities or bargaining problems and achieved the efficient input substitution or the efficient level of trade between the upstream and downstream firms.

Most importantly, I found the evidence of the efficiency effect on the price. The average treatment effect on the price per network is negative and significant at 5 % level except in the case of $M = 4$. Although both groups seem to decrease the price per network during the sample periods, such a decrease is about 13.5 to 20.8 cents more in the merged markets. This difference is quite significant if we consider the fact that the median license fee is about 10 cents.[22]

Other interesting results are that the treatment effects on non-Turner networks owned by Time Warner were positively significant. Specifically, Comedy Central and Court TV increased their market shares about 14-16% and 9-11% more in the merged markets than in the non-merged markets.

The average treatment effect on the penetration rates is not significant although the sign supports the efficiency theories. One explanation for this insignificant result is that vertical integration simply has little effect on them. This may be particularly the case of basic service, which we examine here, because the service is a bundle of affiliated and non-affiliated networks.

5.2 Foreclosure effect

If vertical integration is to facilitate monopolization in the upstream market, then integrated operators will tend to exclude program services, particularly those that directly rival their upstream affiliates. Here, I examine the average treatment effect on the rival networks of Turner Broadcasting. I chose two or three networks in each of four genre: news, children, movie and entertainment, as rivals of *CNN, Cartoon Network, TCM* and *TNT/TBS* respectively. The last category of Table 5, 6 shows the average treatment effect on the rival networks. *CNBC, Fox News Channel, MSNBC* are news channels, *the Disney Channel* and *Nickelodeon* are children channels, *AMC*, *the Independent film channel* and *the Sci Fi Channel* are movie channels, and *Fox Family Channel, FX Network*, and *Odyssey Network* are entertainment channels.

As can be seen in the table, for each genre except news, I found the negative and significant effect on at least one rival network. Specifically, I found that the increase in the market share of the Disney Channel is about 10 to 11.5 percent lower in the merged markets than in non-merged markets with the low significant level. The increases in the market shares of AMC, FX and Odyssey are, respectively, 12.0 to 15.1 percent, 5.9 to 8.4 percent, and

5.2 to 13.0 percent lower in the merged markets. This might suggest that the merged firm practiced foreclosure of these networks.

In addition, I found a positive and significant average treatment effect on the Sci Fi Channel. The reason behind this positive effect may be that this network is actually somewhat related to the new merged firm, Turner-Time-Warner. The Sci Fi channel is partially owned by Liberty Media, which is owned by TCI that owns the share of the Turner-Time-Warner. Hence, similar to the cases of Comedy Central and Court TV, the merged systems seem to favor the networks that are even partially affiliated with the new Turner-Time-Warner, even if they are rival networks of Turner Broadcasting.

5.3 Discussion

The above results show that Time-Warner significantly favors its affiliated networks. Because more downstream markets owned by Time-Warner carried Turner Broadcasting's networks after the merger, the networks of Turner Broadcasting could expand the national market shares significantly. The merger must have benefitted Turner Broadcasting significantly. Even TBS Superstation, which already had 70.4% of the national television household market share in 1995 (Kagan Associate Inc. 2000 pp37), was additionally carried and thereby increased the market share due to the merger. These carriage results support the efficiency theories of vertical integration. The neoclassical efficiency theory shows integration eliminates vertical externalities and promotes efficient input substitution while the organizational economics based efficiency theory argues integration eliminates contractual hazards and thus brings forth the efficient level of trade. Consequently, the trade between the integrated upstream and downstream firms increases. Our results are consistent with this prediction.

The neoclassical efficiency theory predicts lower price and higher sales in downstream markets. As consistent with this theory, I found that the merged markets decreased the price per network more than the non-merged markets. This is a new evidence of the efficiency effect because the previous studies did not find such an effect on prices. Unlike Waterman and Weiss (1996), however, I did not find the positive effect on the penetration rates. This may be because the service studies here is a bundle of affiliated and non-affiliated networks and hence, if the proportion of the integrated networks in the basic services is small, the efficiency gains from them may have little impact on the final sales.[23]

In order to see the network level effect, I also estimated the average treatment effect on the number of subscribers for each network rather than basic services. That is, if a Turner's network is carried in a basic service, then I take the number of subscribers who only subscribe the basic service, and if it is carried in an extended basic service, then I count the number of subscribers in the extended service. This gives me the number of subscribers who actually have access to this network, rather than the number of total subscribers who subscribe to a basic and extended basic services as in the earlier section. If the merger promoted the efficiency for the network, for example, by promoting the downstream's service, then we should see the number of subscribers who have access to the network increases in such markets. The results using 10 matches are shown in Table 7. From this estimation, I found positive and significant average treatment effects of the merger on the number of subscribers who have access to Cartoon Network and Turner Classic Movies. This may suggest that vertical integration led to increased downstream sales effort associated with the two relatively new networks, Cartoon Network and Turner Classic Movies.

Both Chipty (2001) and Waterman and Weiss (1996) found the evidence of foreclosure.

That is, they showed that the vertically integrated cable systems carry rival networks less than the non-integrated cable systems do, as an evidence of exclusionary behavior. Here, I also found the negative treatment effects on several rival networks. However, it may be questionable to describe this phenomenon as an exclusionary behavior because the rival networks seem to have increased their market shares in both merged and non- merged markets. The negative treatment effects on the rival networks merely show that such increases in the market shares are lower in the merged markets. So it is unlikely that this phenomenon is an exclusionary behavior for the purpose of monopolization.

Rather than an exclusionary behavior, the negative treatment effects on the rival networks may be explained as the results from the substitution effects. That is, as in the efficiency theory, the integrated downstream markets substitute more toward the affiliated inputs. Hence "other potential suppliers are in some sense "foreclosed" from providing those input supplied to the vertically integrated firm" (Joskow 2005). In fact, Joskow states that it is important to distinguish between "a naive view" of market foreclosure that is sometimes associated with vertical integration and the issues that arise as a result of the strategic use of vertical integration to soften competition to raise prices. The phenomena seen here seems to fit the former case.

Even if this was an anti-competitive strategic practice, the foreclosure effect, if any, was dominated by the efficiency effect. Along with the results of Chipty (2001) and Waterman and Weiss (1996), the findings in this study suggests there is little evidence that consumers are harmed by the vertical integration.

Either these previous studies or my study, however, examine the effect of foreclosure as a barrier to entry and hence, it may be too early to conclude that foreclosure does not have a negative effect in this industry. By favoring its own children network, Cartoon Network, and excluding its rival networks, the merged firm might have formed a barrier for new children networks to enter in Time Warner's downstream markets. In fact, my dataset shows that no Time Warner markets carry a new children network, *Toon Disney* (that started in 1998) while 31 non-integrated markets carry it with 513,389 subscribers during the examined period.[24] Furthermore, FCC (2001) states that only a few of the 72 planned network services counted in the *1999 Report* have been actually launched during the past year and are now operating, while others have been aborted for various reasons. Such a fact may suggest the difficulty to enter and it may be related to the increased amount of integration in this industry. Further research should be conducted to examine this issue.

6 Conclusion

In this study, I examine the impact of the vertical merger between Turner Broadcasting and Time Warner on the outcomes of the downstream market. Because selection on whether or not participating in this merger is nonrandom, I adopt a non-experimental estimation strategy, the bias-adjusted matching method recently developed by Abadie and Imbens (2002). I modified it by incorporating the difference-in-differences method to mitigate the selection on unobservables problem. For the matching variables, I use nine variables in the participation equation and four variables in the outcome equations. The variables in the participation equation determine whether a system is owned by Time Warner in 1995 (and thus the system is involved with the merger with Turner Broadcasting) and include owner's size, population density, number of homes covered locally, channel capacity, and miles of plant all evaluated at the period of pre-treatment. The variables in the outcome equations are the differenced demand factors that are exogenous to the merger, and include the changes in income, percentage of young, percentage of non-white, and population density.

The results show that there was a significant efficiency gain from the merger. The merged systems were found to carry Turner's networks more frequently than the non-merged systems and there was a larger price decrease in the merged markets. Furthermore, the two new networks, Cartoon Network and Turner Classic Movies, increased their subscribers more in the merged markets.

There is also evidence that the increase in the market shares of the rival networks are lower in the merged markets. Such negative treatment effects, however, should be explained as the results of the substitution effects rather than the exclusionary behavior of foreclosure.

A Estimator

A.1 Evaluation problem

Let Y_{kit} be individual i's outcome measure in period t if she receives treatment W where $W = 0, 1$. The individual treatment impact of moving from treatment state 0 to 1 with outcome measured at time t is

$$\triangle_{it}(0,1) = Y_{1it} - Y_{0it}.$$

If the outcome of interest is price, then the individual impact measures the price increase (or decrease) at period t for a cable system i when the system involved the merger. Because each individual can only receive one and only one treatment in each time period, one of the two terms in this individual treatment impact is not observed. Missing data is a fundamental problem in any impact evaluation study.

Instead of computing the impact for every single individual, a more popular impact parameter of interest measures the average treatment effect on the treated (ATT), which represents the average effect on the outcome for individuals who receive treatment W :

$$\begin{aligned} \tau_t^{TT}(X_{t'} &= x_{t'}) = E(Y_{1t} - Y_{0t}|\, X_{t'} = x_{t'}, W = 1) \\ &= E(Y_{1t}|\, X_{t'} = x_{t'}, W = 1) - E(Y_{0t}|\, X_{t'} = x_{t'}, W = 1) \qquad (1) \end{aligned}$$

where X denotes some observed characteristics with t' denoting the pre-program period and t denoting the post-program period. The second term in the above equation is not observed, since it is the counter factual average outcome in the treatment state $W = 0$ for the individuals who receive treatment $W = 1$. Matching methods provide a way to solve this missing data problem by imputing the counter factual mean using the outcomes of a “matched-comparison” group.

An average version of equation (1), which integrates over the region of support of $X_{t'}$, $S_{X_{t'}}$, can be written as

$$\begin{aligned} \tau_t^{TT}(X_{t'}) &= E(Y_{1t} - Y_{0t}|\, X_{t'}, W = 1) \\ &= \int_{S_X} E(Y_{1t} - Y_{0t}|\, X_{t'} = x_{t'}, W = 1) f_x(X_{t'} = x_{t'}|\, W = 1) dx_{t'} \end{aligned}$$

where $f_x(X_{t'} = x_{t'}|\, W = 1)$ is the conditional density of $X_{t'}$, conditional on receiving treatment W.

The data on the cable systems which were involved in the merger identifies the average treated state outcome measured after the treatment, $E(Y_{1t}|\, X_{t'}, W = 1)$, whereas the data on the cable systems which were not involved the merger gives a direct estimates of $E(Y_{0t}|\, X_{t'}, W = 0)$. Two assumptions are required for matching estimators to identify the parameter of interest, the average treatment effect on the treated (ATT).

Assumption 1: Conditional mean independence

$$E(Y_{0t}|\, X_{t'}, W = 1) = E(Y_{0t}|\, X_{t'}, W = 0) = E(Y_{0t}|\, X) \tag{2}$$

Assumption 2: Existence of comparable non-participants:

$$\Pr(W = 1|\, X_t) < 1. \tag{3}$$

The conditional mean independence condition assumes that the mean outcomes without participation between the treatment participants and the non-participants are equal, conditioning on a vector of variables $X_{t'}$. In other words, conditioning on $X_{t'}$, the mean outcome of the non-participants can be used to impute the mean non-participating outcome for the participants. Assumption 2 simply rules out the possibility of an empty set of nonparticipants at all values of $X_{t'}$. These are weak assumptions for identifying 1. Nonetheless, if unobservables are important determinants of program participation, then the mean independence assumption would not hold. With these two assumptions, the average impact of treatment on the treated can be identified:

$$\begin{aligned} \tau_t^{TT}(X_{t'}) &= E(Y_{1t} - Y_{0t}|\, X_{t'}, W = 1) \\ &= E(Y_{1t}|\, X_{t'}, W = 1) - E(Y_{0t}|\, X_{t'}, W = 1) \\ &= E(Y_{1t}|\, X_{t'}, W = 1) - E(Y_{0t}|\, X_{t'}, W = 0) \end{aligned}$$

where the third equality follows from assumption 2.

A.2 Matching Estimator

If the set of matching variables $X_{t'}$ is large, matching on them becomes problematic because many cells are left empty. The matching estimators recently developed impute the missing potential outcome by using average outcomes for individuals with "similar" values for the covariates. For simplicity, I eliminate the time subscript t in this and following subsections.

Let $\|x\|_V = (x'Vx)^{1/2}$ be the vector norm with positive definite weight matrix V. Define $\|z - x\|_V$ to be the distance between the vectors x and z. Let $d_M(i)$ be the distance from the covariates for unit i, X_i, to the M_{th} nearest match with the opposite treatment. Allowing for the possibility of ties, this is the distance such that strictly fewer than M units are closer to unit i than $d_M(i)$, and at least M units are as close as $d_M(i)$. Formally, $d_M(i) > 0$ is the real number satisfying:

$$\sum_{l:W_i=1-W_i} 1\{\|X_l - X_i\|_V < d_M(i)\} < M \textbf{ and } \sum_{l:W_i=1-W_i} 1\{\|X_l - X_i\|_V \leq d_M(i)\} \geq M,$$

where $1\{.\}$ is the indicator function, equal to one if the expression in brackets is true and zero otherwise.

Let $\vartheta_M(i)$ denote the set of indices for the matches for unit i that are at least as close as the Mth match:

$$\vartheta_M(i) = \{\, l = 1, ..., N|\, W_l = 1 - W_i, \|X_l - X_i\|_V \leq d_M(i)\}.$$

If there are no ties the number of elements in $\vartheta_M(i)$ is M. In general it may be larger. Let the number of elements of $\vartheta_M(i)$ be denoted by $\#\vartheta_M(i)$. Finally, let $K_M(i)$ denote the sum of the weights unit i has as a match for other units, and $K_M'(i)$ the sum of the squared weights in the matches:

$$K_M(i) = \sum_{l=1}^{N} 1\{i \in \vartheta_M(l)\} \cdot \frac{1}{\#\vartheta_M(l)}.$$

$$K_M'(i) = \sum_{l=1}^{N} 1\{i \in \vartheta_M(l)\} \cdot \left(\frac{1}{\#\vartheta_M(l)}\right)^2.$$

Not that $\sum_i K_M(i) = N$, $\sum_{i:W_i=1} K_M(i) = N_0$ and $\sum_{i:W_i=0} K_M(i) = N_1$ where N_0 and N_1 are the number of observations without and with the treatment respectively.

A.2.1 Simple matching estimator

The simple matching estimator uses the following approach to estimate the potential outcome. For each individual i with the treatment $W = 1$, there are two potential outcomes, one is observed and the other is not. The unobserved outcome is the would be outcome of individual i when it were with the treatment $W = 0$, or $Y_i(0)|\, W_i = 1$. Let $\hat{Y}_i(0)$ denote the estimate for this would be outcome. It is estimated by averaging the outcomes of the other most similar individuals who did choose this outcome:

$$\hat{Y}_i(0) = \frac{1}{\#\vartheta_M(i)} \sum_{l \in \vartheta_M(i)} Y_l$$

The simple matching estimator for the average treatment effect for the treated is

$$\hat{\tau}_M^{TT} = \frac{1}{N_1} \sum_{i:W_i=1} (Y_i - \hat{Y}_i(0)) = \frac{1}{N_1} \sum_{i=1}^{N} (W_i - (1 - W_i) \cdot K_M(i)) Y_i.$$

A.2.2 Bias-adjusted matching estimator

The simple matching estimator will be biased in finite samples when the matching is not exact. Abadie and Imbens (2002) show that with k continuous covariates the estimator will have a term corresponding to the matching discrepancies (the difference in covariates between matched units and their matches) that will be of the order $O_p(N^{-1/k})$. In practice one may therefore attempt to remove some of this bias term that remains after the matching. The bias-adjusted matching estimator of Abadie and Imbens adjusts the difference within the matches for the differences in their covariate values. The adjustment is based on an estimate of the regression function $\mu_0(x) = E[Y(0)|\, X = x]$. Following Adabie and Imbens (2002) I approximate this regression function by linear function and estimate it using least squares on the matched observations:

$$\hat{\mu}_0(x) = \hat{\beta}_{00} + \hat{\beta}_{01}' x,$$

where

$$(\hat{\beta}_{00}, \hat{\beta}_{01}) = \arg\min \sum_{i:W_i=0} K_M(i) \cdot (Y_i - \beta_{00} - \beta_{01}' X_i)^2.$$

Given the estimated regression functions, they predict the missing potential outcomes as;

$$\tilde{Y}_i(0) = \frac{1}{\#\vartheta_M(i)} \sum_{l \in \vartheta_M(i)} (Y_l + \hat{\mu}_0(X_i) - \hat{\mu}_0(X_l)).$$

The bias -adjusted matching estimator for ATT is

$$\hat{\tau}_M^{TT,bc} = \frac{1}{N_1} \sum_{i:W_i=1} (Y_i - \tilde{Y}_i(0)).$$

A.2.3 Bias-adjusted matching estimator with difference-in-differences

As I have a richer data that contains the pre and post treatment outcomes, I also adopt the difference-in-difference method of Heckman et. al (1997, 1998). An equivalent version of (2) for the difference-in-difference matching estimator can be expressed as

$$E(Y_{0t} - Y_{0t'} | X_t, W = 1) = E(Y_{0t} - Y_{0t'} | X_{t'}, W = 0) = E(Y_{0t} - Y_{0t'} | X).$$

The general form of this estimator can be written as

$$\hat{\tau}_M^{TT,bc,did} = \frac{1}{N_1} \sum_{i:W_i=1} (\triangle Y_i - \triangle \tilde{Y}_i(0)).$$

where $\triangle$ expresses the difference between the post and the pre treatment values.

B Structural Model

This model follows Waterman and Weiss (1996) with modification. While they focused on pay services of the cable systems, I focus on basic services as all of the networks involved in the merger I will examine are basic networks. The model for basic services may be more complicated because of the fact that they are bundles of many networks.

The model consists of three agents, cable systems $i = 1...I$, consumers in systems' franchise areas and cable networks $j = 1...J$. In this industry, cable networks are upstream firms while cable systems are downstream firms. We also consider the existence of cable operators and cable programming firms. We use the term “cable systems” for individual cable distributors in cable franchise markets. Most of cable systems are monopolies in their franchise areas. We use the term “cable operators” for the firms who own multiple cable systems (MSOs). For example, Time Warner is a cable operator who owns many cable systems. Similarly, we use the term “cable networks” for individual cable TV channels and the term “cable programming firms” for the firms who own many networks. For example, Turner Broadcasting is a cable programming firm who own many cable networks such as CNN.

We assume that each agent's behavior at time t does not affect the same agent behavior at different time t'. That is, we do not consider the dynamic effect and the solutions to the agents from optimization are same for every t. For notational simplicity, I omit the time subscript in this sub-section.

Profit of cable system i when they are not integrated is given as

$$\pi_i = p_i s_i - C(s_i, \{NET_{ij} : j = 1...J\}, \vec{w}_{Bi}, T_i, c_i), \tag{4}$$

where p_i is price, s_i is subscribership (or demand), C is the cost function which depends on input cost $\vec{w}_B$, the network offering decision NET_{ij} ($NET_{ij} = 1$ if if the system i carry the network j in its basic service), and other cost T. Other costs are those such as service and line maintenance, and depend on various technological factors such as the size and density of the local service area. We assume that the input cost $\vec{w}_B$ is solely consists of the license fee that systems pay to networks and thus, $\vec{w}_{Bi} = \{w_{ij} : j = 1, ...J\}$ where w_{ij} is the license fee (per subscriber per month) of network j for system i. Note that the cost minimization problem involves license fee of all the networks, i.e. even if system i does not carry network j, w_{ij} is involved in the cost function. c_i contains unobserved firm (local system) characteristics such as managerial quality or structure, that can be viewed as being (roughly) constant over the period.

The subscribership or consumers' demand in system i's franchise area is given by (this should be driven by consumers' utility maximization)

$$s_i = S(p_i, q_{Bi}, u_{Bi}, G_i, D_i), \tag{5}$$

where q_{Bi} is observed quality/characteristics of system i's basic service, u_{Bi} represents unobserved (by econometrician) systematic consumers' preference to the system i's basic service, constant over the period. For example, the availability of the satellite TV depends on geographical characteristic, is constant overtime, and affects the demand on the cable TV. G_i is a vector of demographic variables and D_i is a vector of other demand factors such as prices and availability of competing entertainment services.

The profit of the network j is given as

$$\pi_j = \sum_i^I w_{ij} NET_{ij} s_i - F_j \tag{6}$$

where Fj is the fixed cost for production of network j. We ignore the marginal cost for network production because it is almost negligible compared with the fixed cost.

Regarding the vertical structure, we assume that the upstream firms (cable networks) set their price first followed by the downstream firms (cable systems). We also assume that the cable operator has decision right to determine the behavior of systems they own. Then the maximization problem of a cable operator which owns multiple systems $i' = 1...I'$ is

$$\begin{aligned}
\max_{\{p_i, NET_i, ; i'=1...I'\}} \sum_{i'}^{I'} \pi_i &= p_i s_i - C(s_i, NET_i, \vec{w}_{Bi}, T_i, c_i) \\
s.t.\ s_i &= S(p_i, q_{Bi}, u_{Bi}, G_i, D_i) \\
\{w_{i'j} : j &= 1...J, i' = 1...I'\} \\
\sum_j NET_{ij} &\leq Z_i
\end{aligned} \tag{7}$$

where Z_i represents the carriage capacity of the system i and $NET_i = \{NET_{ij} : j = 1...J\}$.

But in this case, as a cable system's behavior does not affect another system's profit, each individual cable system maximizes its profit in its franchise area regardless of whether it is owned by MSO. That is,

$$\max_{\{p_i, NET_i, ; i'=1...I'\}} \sum_{i'}^{I'} \pi_i = \sum_{i'}^{I'} \max_{\{p_i, NET_i : i'=1...I'\}} \pi_i.$$

The solutions of this problem lead us to have

$$p_i^* = p^*(T_i, \vec{w}_{Bi}, q_{Bi}, u_{Bi}, D_i, G_i, Z_i, c_i) \tag{8}$$

$$NET_{ij}^* = N^*(T_i, \vec{w}_{Bi}, q_{Bi}, u_{Bi}, D_i, G_i, Z_i, c_i) \tag{9}$$

Then the demand equation (5) becomes

$$s_i^* = S^*(T_i, \vec{w}_{Bi}, q_{Bi}, u_{Bi}, D_i, G_i, Z_i, c_i). \tag{10}$$

Given these equations, cable networks maximize their profit (6) with respect to w_{ijt} (for simplicity, we assume that a cable programming firm maximizes profit of each network it owns independently). The solution can be expressed as[25]

$$w_{ij}^* = w(\{T_i, w_{Bi}, q_{Bi}, u_{Bi}, D_i, G_i, Z_i, c_i : j = 1...J, i = 1...I\}, F_j). \tag{11}$$

As seen, the franchise fee of network j for system i depends on not only its franchise fee for all other system $i' \neq i$, but also the franchise fee of all the other network $j' \neq j$ for all the cable systems.

Solving (8), (9), (10), and (11), we will have

$$
\begin{aligned}
p_i^* &= p^*(\{T_i, q_{Bi}, u_B, D_i, G_i, Z_i, c_i, F_j : j = 1...J, i = 1...I\}) && (12) \\
NET_{ij}^* &= N^*(\{T_i, q_{Bi}, u_{Bi}, D_i, G_i, Z_i, c_i, F_j : j = 1...J, i = 1...I\}) && (13) \\
s_i^* &= S^*(\{T_i, q_{Bi}, u_{Bi}, D_i, G_i, Z_i, c_i, F_j : j = 1...J, i = 1...I\}) && (14) \\
w_{ij}^* &= w(\{T_i, q_{Bi}, u_{Bi}, D_i, G_i, Z_i, c_i, F_j : j = 1...J, i = 1...I\}). && (15)
\end{aligned}
$$

As our primary aim is to see the *net* effect of vertical integration on these endogenous variables and not to estimate parameters of supply-demand functions, we can achieve our goal by estimation on these reduced forms. Because we do not have license fee for each system w_{ij} in our data, we will actually estimate the equations of price (12), network offering (13), and subscriber (14).

References

[1] Abadie A. and G. Imbens, "Simple and Bias-Corrected Matching Estimators for Average Treatment Effects", NBER Technical Working Paper 283.

[2] Abadie A., D. Drukker, J. Herr, and G. Imbens (2001), "Implementing Matching Estimators for Average Treatment Effects in Stata", The Stat Journal 1, 1-18.

[3] Aghion, P and P. Bolton (1987), "Contracts as a Barrier to Entry," *American Economic Review*, 77, 388-401

[4] Chipty T. (2001), "Vertical Integration, Market Foreclosure, and Consumer Welfare in the Cable Television Industry", *American Economic Review* 91 428-451.

[5] Federal Communications Commission (2001) "Seventh Annual Report".

[6] Gilbert, R. and J. Hastings, (2001), "Vertical Integration in Gasoline Supply: An Empirical Test of Raising Rivals' Costs", Working Paper E01-302, Department of Economics, University of California, Berkeley, CA.

[7] Hart, O. and J. Tirole (1990), "Vertical Integration and Market Foreclosure, " *Brookings Papers on Economic Activity*, Special Issue: 205-276.

[8] Heckman J., Ichimura, H. and P. Todd, (1997), "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", *Review of Economic Studies* 64, 605-654.

[9] Heckman J., Ichimura, H. and P. Todd, (1998), "Matching As An Econometric Evaluation Estimator", *Review of Economic Studies* 65, 261-294.

[10] Imbens G. (2004), "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review", *Review of Economics and Statistics* 86, 4-29.

[11] Joskow P. (2005), *Handbook of New Institutional Economics*, Kluwer, forthcoming.

[12] Kagan Associates Inc. (2000), *Kagan's Economics of Basic Cable Networks.*

[13] Moffit R. (2004), "Introduction to the Symposium on the Econometrics of Matching", *Review of Economics and Statistics* 86, 1-3.

[14] Ordover, J., S. Salop and G. Saloner (1990), "Equilibrium Vertical Foreclosure, " *American Economic Review* 80, 127-42.

[15] Tirole J. (1998) *The Theory of Industrial Organization*, Cambridge, MA, MIT Press.

[16] Vita, M. (2000) "Regulatory Restrictions on Vertical Integration and Control: The Competitive Impact of Gasoline Divorcement Policies," *Journal of Regulatory Economics* 18, 217-33.

[17] Warren Publishing (1996, 2001), "Television and Cable Factbook".

[18] Waterman D. and A.A. Weiss (1996), "The Effects of Vertical Integration Between Cable Television Systems and Pay Cable Networks," *Journal of Econometrics* 72, 357-395.

[19] Williamson, O. (1975), *Markets and Hierarchies: Analysis and Antitrust Implications,* New York, Free Press.

[20] Williamson, O. (1983), "Credible Commitments: Using Hostages to Support Exchange, " *American Economic Review* 73, 519-540.

[21] Williamson, O. (1985), *The Economic Institutions of Capitalism*, New York, Free Press.

[22] Williamson, O. (2003), "On the Transaction Cost Determinants of Vertical Integration," *Journal of Law, Economics and Organization*, 19, 1-23.

[23] Wooldridge J. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press.

	Mean	S.D.	Treated	Control	DID	
Change in Turner Number	0.581	1.089	0.912	0.467	0.445	***
Change in Non-Turner Number	14.654	10.458	16.976	13.826	3.150	***
Change in Fee per Network	-0.425	1.090	-0.383	-0.440	0.565	
Change in Penetration Rate	0.457	6.178	0.510	0.440	0.070	
Change in Prob(TBS)	0.058	0.350	0.106	0.041	0.065	**
Change in Prob(Cartoon)	0.182	0.398	0.315	0.137	0.179	***
Change in Prob(CNN)	0.006	0.307	-0.006	0.010	-0.016	
Change in Prob(TCM)	0.176	0.391	0.255	0.149	0.106	***
Change in Prob(TNT)	0.101	0.362	0.081	0.108	-0.026	
Change in Prob(Comedy)	0.315	0.468	0.445	0.271	0.175	***
Change in Prob(CourtTV)	0.081	0.302	0.194	0.042	0.152	***
Change in Prob(CNBC)	0.029	0.441	0.055	0.020	0.035	
Change in Prob(Fox News)	0.109	0.314	0.070	0.088	0.082	***
Change in Prob(MSNBC)	0.166	0.376	0.276	0.128	0.148	***
Change in Prob(Disney)	0.229	0.426	0.064	0.285	-0.222	***
Change in Prob(Nickelodeon)	0.221	0.437	0.239	0.215	0.024	
Change in Prob(AMC)	0.568	0.509	0.624	0.549	0.075	***
Change in Prob(Independent)	0.002	0.039	0.003	0.001	0.002	
Change in Prob(Scifi)	0.203	0.423	0.248	0.187	0.061	**
Change in Prob(FoxFamily)	0.043	0.273	0.048	0.042	0.006	
Change in Prob(FX)	0.091	0.293	0.118	0.082	0.036	**
Change in Prob(Odyssey)	0.140	0.349	0.215	0.114	0.101	***

Stars refer to the significance of a t-test on the means.
*=significant at the 10% level, **=significant at the 5% level, ***=significant at the 1% level

Table 1: Dependent variables

Theoretical Construct	Empirical measure	Level
Demographic variables		
Income	Median household income (dollars)	County
Size of TV market	Number of homes covered locally	System
Younger viewership	Percentage of population between age 5 and 15	County
Non-white viewership	Percentage of population non-white	County
System and owner characteristics		
System efficiency	Number of years since franchise began	System
System efficiency	Population density (1992 population per 100- square miles)	County
System size	Number of homes covered locally	System
System size	Channel capacity	System
System size	Miles of plant	System
Owner characteristics	Total number of homes covered by an owner	Owner

Table 2: Matching variable definitions

The variables in the participation equation (Z)	
Owner size	Total number of homes covered by an owner in 1995
Income	Median household income in 1992
Young	Percentage of population between age 5 and 15 in 1992
Non White	Percentage of population non-white in 1992
Density	Population density in 1992
Age	Number of years since franchise began in 1995
Home passed	Number of homes covered locally in 1995
Capacity	Channel capacity in 1995
Plant	Miles of plant in 1995
The variables in the outcome equations (V)	
Change in Income	Change from 1992 to 2000
Change in Young	Change from 1992 to 2000
Change in Non White	Change from 1992 to 2000
Change in Density	Change from 1992 to 2000

Table 3: Matching variables

	Mean	**S.D.**	**Treatment**	**Control**	
Ownersize	1,191,758	1,806,763	2,761,317	668,070	***
Income	32,238.02	7,133.73	33,116.51	31,944.62	***
Young	18.33	2.22	18.39	18.31	
Non white	11.82	12.34	12.76	11.51	***
Density	391.22	1938.97	473.86	363.62	**
Age	17.73	9.66	20.83	16.60	***
Home passed	18,242.24	63,228.11	30,712.34	13,997.72	***
Capacity	41.86	14.71	42.82	41.52	***
Plant	187.47	411.28	308.92	146.95	***
Change in income	3304.23	2487.23	3283.59	3311.12	
Change in young	0.03	1.24	0.11	0.00	***
Change in non white	0.71	2.80	0.43	0.80	***
Change in density	36.12	409.49	70.29	24.71	***
Number of markets			348	1,043	

Stars refer to the significance of a t-test on the means. *=significant at the 10% level, **=significant at the 5% level, ***=significant at the 1% level.

Table 4: Summary Statistics of Matching Variables

	$M=4$		$M=6$		$M=8$		$M=10$	
Change in	0.289	**	0.385	***	0.427	***	0.444	***
Turner Number	(0.140)		(0.146)		(0.140)		(0.145)	
Change in	-0.892		-0.763		-0.759		-0.754	
Non-Turner Number	(1.224)		(1.248)		(1.259)		(1.241)	
Change in	-0.134		-0.167	**	-0.193	**	-0.208	**
Fee per Network	(0.094)		(0.082)		(0.080)		(0.083)	
Change in	0.455		0.427		0.432		0.452	
Penetration Rate	(0.596)		(0.574)		(0.562)		(0.629)	
Change in	0.113	**	0.118	**	0.127	***	0.125	**
Prob(TBS)	(0.048)		(0.046)		(0.046)		(0.046)	
Change in	0.043		0.058		0.044		0.041	
Prob(Cartoon)	(0.055)		(0.058)		(0.053)		(0.052)	
Change in	-0.083	**	-0.069	**	-0.056	*	-0.055	*
Prob(CNN)	(0.035)		(0.032)		(0.031)		(0.031)	
Change in	0.093	**	0.126	**	0.137	***	0.145	***
Prob(TCM)	(0.047)		(0.049)		(0.050)		(0.052)	
Change in	-0.005		0.021		0.043		0.052	
Prob(TNT)	(0.044)		(0.044)		(0.043)		(0.052)	
Change in	0.141	***	0.162	***	0.154	***	0.149	***
Prob(Comedy)	(0.050)		(0.050)		(0.047)		(0.045)	
Change in	0.088	**	0.108	***	0.107	***	0.101	***
Prob(Court TV)	(0.038)		(0.038)		(0.037)		(0.037)	

Note: Weighting matrix: inverse matrix
Matching variables: ownersize, income, young, non-white, density, age, homepassed, capacity,plant, change in income, change in young, change in non-white, change in density
Bias adjusted variables: all the matching variables
Stars refer to the significance leve *=10% leve, **=5% level, ***=1% level

Table 5: Average Treatment Effects for the Treated

	$M = 4$	$M = 6$	$M = 8$	$M = 10$
Change in	-0.019	0.019	0.031	0.039
Prob(CNBC)	(0.047)	(0.047)	(0.045)	(0.044)
Change in	-0.037	-0.063	-0.084 **	-0.089 **
Prob(Fox News)	(0.044)	(0.043)	(0.041)	(0.040)
Change in	0.038	0.023	0.004	-0.003
Prob(MSNBC)	(0.043)	(0.041)	(0.042)	(0.041)
Change in	-0.115 **	-0.100 *	-0.102 *	-0.104 *
Prob(Disney)	(0.052)	(0.055)	(0.059)	(0.060)
Change in	-0.019	-0.017	0.000	-0.003
Prob(Nickelodeon)	(0.049)	(0.050)	(0.046)	(0.045)
Change in	-0.120 **	-0.126 ***	-0.141 ***	-0.151 ***
Prob(AMC)	(0.051)	(0.047)	(0.047)	(0.047)
Change in	0.003	0.003	0.003	0.003
Prob(Independent)	(0.006)	(0.006)	(0.006)	(0.005)
Change in	0.048	0.085 *	0.085 *	0.090 **
Prob(Scifi)	(0.047)	(0.046)	(0.045)	(0.044)
Change in	0.041	0.042	0.055	0.055
Prob(FoxFamily)	(0.036)	(0.036)	(0.035)	(0.036)
Change in	-0.059 *	-0.065 *	-0.079 **	-0.084 **
Prob(FX)	(0.035)	(0.035)	(0.036)	(0.036)
Change in	-0.052	-0.100 **	-0.121 ***	-0.130 ***
Prob(Odyssey)	(0.044)	(0.045)	(0.047)	(0.004)

Note: Weighting matrix: inverse matrix
Matching variables: ownersize, income, young, non-white, density, age, homepassed, capacity,plant, change in income, change in young, change in non-white, change in density
Bias adjusted variables: all the matching variables
Stars refer to the significance leve *=10% leve, **=5% level, ***=1% level

Table 6: (Continued)

	(M=10)
Change in sub(TBS)	1572.57
	(6159.30)
Change in sub(Toon)	10879.42 *
	(5789.94)
Change in sub(CNN)	5096.34
	(3397.37)
Change in sub(TCM)	11318.09 **
	(4667.34)
Change in sub(TNT)	2789.32
	(4513.22)

Table 7: Average Treatment Effects for the Treated on the number of subscribers for each network