The Dissertation Committee for Changseung Yoo
certifies that this is the approved version of the following dissertation:

Exploring Horizontal and Vertical Interactions within an Online Advertising Supply Chain

Committee:

Genaro J. Gutiérrez, Supervisor

Anitesh Barua, Co-Supervisor

Garrett Sonnier

Ioannis Stamatopoulos

Exploring Horizontal and Vertical Interactions within an Online Advertising Supply Chain

by

Changseung Yoo

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

December 2018

Dedicated to my wife Haejung.

Acknowledgments

A lot of people helped me, wittingly or unwittingly, through my PhD journey. I am forever indebted to each and every person involved. First and foremost, I am grateful for the love and emotional support from my wife, Haejung Yi, and my two adorable angels, Chaewoo and Jiwoo. I would not have made it through without them. Other family members provided substantial assistance as well. My mother, father, and brother were always there for me, and my in-laws have pulled for me each step of the way. All of them have been more important to me during my studies than they realize.

I also sincerely thank my two advisors, Genaro J. Gutiérrez and Anitesh Barua, who have been extremely patient, encouraging, and supportive throughout my PhD career. They have helped me broaden my intellectual horizon and transform the way I view the world. Every single idea in this document we have brainstormed together, formalized together, and reiterated numerous times.

I am deeply thankful also to my friends here and abroad. Thank you for not allowing me to give up when I am frustrated. Thank you for making me laugh even when I am tense or discouraged for no apparent reason. Thank you (especially) for your company, physical or virtual, when I absolutely needed them.

Exploring Horizontal and Vertical Interactions within an Online Advertising Supply Chain

Publication No. ______________

Changseung Yoo, Ph.D.
The University of Texas at Austin, 2018

Supervisor : Genaro J. Gutiérrez
Co-Supervisor : Anitesh Barua

This dissertation consists of highly related three essays on estimating and exploiting the effect of complementarities or synergies between advertising channels within an online advertising ecosystem. The first essay studies how the recognition of complementarities between different channels in the supply chain and sharing information among players can lead to dramatic increases in profitability and scale of operations. Using a naturally occurring experiment, the second essay investigates the impact of decision making structure (centralized vs. decentralized decision-making) on the decision maker's ability to exploit the synergy effects between advertising channels. Lastly, the third essay explores through simulation the challenges involved in estimating complementarities from a structural econometric model based on the researcher's knowledge of whether the decision makers actually incorporate such effects in their decisions.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Online advertising is a fascinating research area since it inevitably involves rich data and unique research questions that are of interest in various fields of business, economics, and engineering. During my doctoral studies at McCombs School of Business, I had the privilege of interacting with practitioners from this dynamically evolving industry. Through these interactions, I was not only able to understand the key elements/players of the online advertising ecosystem and business-to-business contracting practices, but I also had the opportunity to learn first hand the day-to-day operations and decision-making practices by the managers in the agency. These valuable experiences were instrumental in developing research ideas and precisely pinpointing my research questions.

Many researchers have investigated the online advertising industry from various angles including, for example, how to increase online advertising effectiveness and how to make the real-time-bidding auction mechanism more efficient. However, my experience with the online advertising agency made me realize that there is a new angle to be explored, that there is a complex underlying structure in which countless entities (e.g., advertisers, ad agencies, publishers, ad aggregators) as well as their interactions are present; that is, there exists a digital *supply chain* of online advertisements, the characteristics of which cannot be adequately explained by traditional supply chain management literature. Given my background and training in supply chain management, I became interested in exploring further the roles that each of the entities play as well as the impact of their decisions and interactions among them on campaign performance and system-wide profit.

In the early phase of interaction with the ad agency, we learned about the company's perceived economic role within the digital supply chain. The agency perceived its role as exploiting economies of scale in media buying (advertising slot aggregator) and lowering search and transaction costs for the advertisers who carry out online ad campaigns through the agency. As a result, the company's decision-making structure centered on specialized groups of managers who ran multiple campaigns on their respective platforms. However, I observed that there is a potentially more important economic rationalization for their role; they can also be information aggregators who leverage this information to exploit synergies among channels increasing significantly system-wide profits. Three distinct but highly related essays spawned from this idea, and this dissertation is comprised of those essays. Below I outline the key ideas of these three papers in more detail.

Essay 1: Strategic Complementarities in an Online Advertising Supply Chain

This paper explores a supply chain of online advertising which comprises of advertisers, agencies, publishers, brokers, and their interactions. Using data from a publicly traded, multinational advertising agency, we explore the intra- and inter-channel interactions in its digital supply chain, and analyze the impact of interactions between channel structures on the agency's decisions and campaign performance. The extant literature, which offers advice on increasing the click-through-rate and/or conversion rate, has largely ignored interactions within the supply chain. Further, the supply chain literature, while vast, has not focused on the domain of online advertising, and has therefore not studied the structure or interactions in such a digital setting. Through our empirical investigation using data on a campaign that lasted 382 days, we find that there are substantial vertical interactions and horizontal synergies, the failure to account for which may lead to significant underspending and suboptimal performance. Incorporating such interactions and synergies in the agency's decision making can potentially increase the overall supply chain profit by 78% over the status quo. In addition, we show that by combining information from multiple channels and structuring contracts appropriately, the agency can help further boost the performance of the supply chain. Our results offer practical guidance for improving the performance of the $239 billion online advertising industry through the analysis of various interactions, information and profit sharing schemes. Our study also shows that the agency, while currently organized by vertical media buying, should instead be re-organized by campaigns in order to monetize the substantial benefits of cross-channel complementarities.

Essay 2: The Effect of Organizational Design on Exploiting Complementarities

Using a naturally occurring experiment, we empirically investigate how a firm's organizational design can affect its ability to exploit synergies among business practices in the context of digital advertising. Specifically, we compare the decisions and performances of two different decision-making structures by managers in an online advertising agency: the centralized mode where a single manager runs an ad campaign on two advertising channels simultaneously, and the decentralized mode where two different managers run the same ad campaign in their respective channels. We find that the centralized mode better recognizes and incorporates complementarities between the two channels into its decision making, while the decentralized mode fails to understand complementarities and thus systematically fails in allocating the right amount of resources in each channel, leading to less overall profit for the agency. Our analysis demonstrates that the decision structure can make a difference in the performance due to more effective exploitation of cross-channel synergies in an online advertising campaign.

Essay 3: On the Challenges of Detecting Complementarities in a Structural Model

In the empirical literature in economics and information systems on testing complementarities, no distinction has been made between (i) the existence of complementarities and (ii) whether decision makers recognize and act on such synergies. To address this gap, we analyze how a decision maker's understanding of complementarities can affect the empirical evidence of complementarity or substitutability within a structural econometric model. We study the decisions of an agency in an online advertising supply chain with substantial prima facie justification of complementarities across its channels of operation. However, from actual observations of its operations, we have inferred that the agency makes decisions without acting on such synergies. We find that the instrumental variables (IV) approach establishes the presence of strong complementarities regardless of whether or not the agency acts upon complementarities; however, two structural models, one in which the researcher mistakenly assumes the decision process to incorporate complementarities, and the other in which the researcher is informed about the naive decisions, fail to extract the interactions, and produce a significant bias in the estimates, whereby they may even appear to be substitutes. Our results imply that unless the researcher is certain about the decision maker acting on complementarities, it is prudent to rely on the IV approach than a structural model to establish the presence of and estimate the magnitude of complementarities. An important caveat, however, is that the IV approach can sometimes be highly inefficient relative to the explicit structural model.

Chapter 2

Strategic Complementarities in an Online Advertising Supply Chain

2.1 Introduction

The global online advertising revenues will reach a staggering $229 billion in 2017, long surpassing spends on cable and broadcast television advertising in 2011 and 2013, respectively (Liu 2016). Along with the rapid rise in spend on online advertising, sophisticated channel structures, pricing mechanisms and technology-driven platforms have emerged to help advertisers broaden their reach, find new customers and ultimately increase online sales. With an ever widening set of choices of online media, advertising formats, and pricing methods, a complex digital supply chain has evolved in this industry. However, the interactions and decisions within the supply chain, which can affect its performance, have not been investigated thoroughly in the extant literature. What is the structure of this supply chain? How are the decisions of various players affected by those of others, and what is their impact on profitability? Are there any synergies across channels within this supply chain? How can campaign performance be improved through information sharing and other means such as profit sharing within the supply chain? These are some of the research questions we address in the study.

An online advertising supply chain consists of publishers, advertising agencies, brokers, and advertising channels. We consider two distinct advertising channels: an advertising *network* and an advertising *exchange* for trading impressions[1] and/or conversions[2] as shown in Figure 2.1. The advertising network is a closed group of specific publishers analogous to a privately traded market, where prices are determined through individual negotiations, while the advertising exchange is a technology-driven digital marketplace that facilitates buying and selling of advertising slots with prices determined by a real-time-bidding (RTB) mechanism. Major ad exchange operators include Google DoubleClick Ad Exchange, Microsoft Ad Exchange, and OpenX. Publishers upload advertisers' creatives on their websites, and agencies help advertisers with tactics to convey their messages, dispatch materials to potential customers through publishers to achieve campaign objectives. Brokers aggregate advertising slots from various publishers and sell them to advertising agencies through the exchange platform. Individual publishers in the advertising network are generally specialized in focus (e.g., sports, news, or entertainment), while the exchange makes a wide range of publishers available, which the agencies may not have found otherwise due to search costs and other constraints.

⟨⟨Figure 2.1 about here⟩⟩

Approximately 64% of 2016 revenues in the U.S. were priced on a

[1] An event that is counted when an ad is displayed on a webpage.

[2] An event that is counted when a customer interacts with an ad and takes a measurable action such as purchasing a product or downloading online content.

performance basis[3], and 35% were priced based on cost per mille (CPM) or impressions (Interactive Advertising Bureau (IAB) 2016 Internet Advertising Revenue Report). While performance-based pricing has been the lead pricing model since 2006, many online advertising agencies use both pricing models concurrently, and may even use hybrid pricing[4] which constitutes approximately 1% of the 2016 revenues. Our study involves both performance and CPM based pricing.

Prior marketing and operations management literature has primarily focused on user interactions with online ads, and has implicitly treated the supply chain of online ad inventory as a black box. Some studies offer advice on how an advertiser can increase the click-through-rate (CTR)[5] and/or conversion rate (CR)[6] of a campaign by investigating the interactions between an advertising format (e.g., paid search, display, etc.) and consumers, while putting less emphasis on potential interactions among the channels themselves (e.g., [20], [45], and [54]). Research has also explored synergies across different formats of online advertising (e.g., [58], [1], and [57]), but has paid less attention to potential synergies between channel structures or between pricing

[3] In performance-based advertising, the advertiser pays only for measurable results such as leads, downloads and sales. Pricing models in performance-based advertising include cost-per-click, cost-per-lead, and cost-per-acquisition.

[4] Hybrid pricing uses a combination of a CPM pricing model and a performance-based pricing model.

[5] CTR is the number of clicks on an ad divided by the number of times the ad is shown, i.e., CTR = clicks/impressions.

[6] The typical industry definition of CR is the number of conversions from an ad divided by the number of clicks on the ad, i.e., CR = conversions/clicks. In this paper, however, we will define CR to be the number of conversions divided by the number of impressions.

models within the advertising supply chain.

Aided by the opportunity to work with a multi-national advertising agency[7], our study takes an early step in analyzing the complete online advertising supply chain, tracing and quantifying the interaction effects between channel structures and their impact on the agency's decision making and profitability. Specifically, we focus on interactions between pricing schemes as well as between the network structure and the exchange platform. Conversions in the two channels may not be independent of each other as the impressions from both channels are targeted toward the same set of customers. The prima facie justification of positive interactions or complementarities is that repeated exposure across multiple websites may entice a consumer to click on an ad, and even buy the advertiser's product, as studied in the literature on interaction between formats (e.g., [58], [1], and [57]). Another rationale is that an agency can learn through historical data about the effectiveness of, say, impressions from websites in the exchange, and apply such knowledge in choosing actions in the network. However, [21], [22], and [23] show that there can be substitution effects across advertising media. For example, [23] find that if highly targeted plain text ads and more visually striking but less targeted ads are used in combination, the two strategies are ineffective due to privacy concerns.

Based on the understanding we developed by observing the operations at the ad agency, we model the decision making processes of a publisher in

[7] With offices in 11 countries worldwide, the company has over 5,000 active customers with over 18,000 live campaigns and over 120 billion impressions per month.

the network and the agency as a Stackelberg game. This model helps derive insights into key drivers of the players' decisions. We then apply such insights into choosing a set of instrumental variables in our simultaneous equations estimation procedure, and show that there are significant synergy effects between channel structures and pricing models. Specifically, we find that the agency's expected profit function is supermodular in the agency's actions in each channel, and that the number of conversions from each channel is supermodular in the channel impressions. We demonstrate via counterfactual analysis that incorporating such complementarities in the agency's decision making process can increase the supply chain's profit by 78% over the status quo. Moreover, information and profit sharing schemes can more than triple the agency's profit, bringing it closer to that of a theoretical benchmark scenario, where the agency and the publisher in the network are an integrated entity. The key managerial implications of our research are that the supply chain is more profitable when the agency shares information and/or profit with the publishers in the network, and that the agency should restructure its organization to facilitate information flow between decision makers in the two channels, and implement cross-training or centralized decision making among managers to fully exploit the effect of complementarities and realize maximum benefits. Thus our study provides the rationale for a new economic role for the agency, which creates value not only through economies of scale in media buying, but also through information and coordination based intermediary activities in the supply chain.

Our study contributes to the nascent literature on synergy effects in online advertising, providing both analytical and empirical support for complementarities between channel structures and between pricing models leading to higher supply chain profits and levels of operation. Quantifying vertical and horizontal interaction effects in the online advertising supply chain is useful for the online ad industry, not only because it can boost profits, but also because it provides a rationale to rethink internal decision structures within the agency as well as information sharing and incentives along the supply chain.

2.2 Literature Review

Since our research is interdisciplinary in nature, we review multiple streams of relevant literature including studies in online advertising effectiveness, synergy effects in online advertising, supply chain coordination and contracts, and empirical studies in supply chain management, which help identify opportunities for analyzing the online advertising supply chain.

2.2.1 Online Advertising Effectiveness

Our research is closely related to the literature on online advertising effectiveness. [33] study the effect of banner advertising on current customers' repurchase probabilities, while [20] and [45] investigate factors driving consumers' search and conversion behavior in the sponsored search advertising domain. [22] show that matching an ad to website content and increasing an ad's obtrusiveness independently increase purchase intent, but the two strategies are ineffective when used concurrently. In a more recent paper, [54] explores the effectiveness of social advertising using data from ads on Facebook.

In contrast to the above list of studies, we study the effectiveness of online ads taking into account potential interactions between ads in a multi-channel context. Furthermore, we focus on the advertising agency's role as a decision maker of campaigns, and study its expected profit maximization problem in a Stackelberg framework. Finally, we devise feasible schemes that can boost the supply chain's profit as a whole, and find a new rationalization for the agency's role as a digital intermediary.

2.2.2 Synergies in Online Advertising

Our study largely draws upon the nascent literature of synergies among multiple modes or formats of online advertising. [58] show that organic and sponsored search advertising have positive interdependence. [1] estimate a hidden Markov model of consumer behavior and find that display and search ads affect customers differently based on their states in the purchasing process. [57] use a mutually exciting point process to show that display advertisements stimulate subsequent visits to other advertising formats and eventual conversions. [25] show that combining web and mobile display ads performs better than when either web or mobile is used in isolation. [29] find that display ads increase search clicks and conversions.

Our work differs from the above literature in two distinct ways. First, prior literature does not consider the larger ecosystem of online advertising. Instead, these studies primarily focus on the interaction between ads on websites and individuals visiting them, and analyze strategies to increase the likelihood of clicks or conversions. By contrast, our model involves the larger supply chain of online ads, starting from the supply side (websites) to advertisers, and entities and channels located between them, which allows us to further investigate different facets of this complex system. Second, the above literature does not study potential interaction or synergies between different pricing models and between different platforms (e.g., network and exchange). In fact, some studies have analyzed when to use a particular pricing scheme to the exclusion of others (e.g., [44], and [27]). Our study demonstrates that these pricing mechanisms and platforms are not mutually exclusive, and that ignoring the interplay between these choices can lead to underspending. Thus, our work complements the extant literature and provides detailed analysis to gain a better understanding of underlying online ad supply chain structures that affect campaign performances.

2.2.3 Supply Chain Coordination and Contracts

A supply chain consists of multiple decision makers possibly having different incentives and information, which may lead to suboptimal supply chain performance. This phenomenon has been widely studied in the economics and operations management literatures; the first concrete analysis is often attributed to [50], who first described the *double marginalization* problem. Ideally, a single decision maker would optimize the supply chain with all information at hand in order to avoid incentive coordination problems. Researchers in operations management refer to this system as the *integrated* supply chain. However, integrating a supply chain may often be impractical due to economic, administrative, and other constraints. As a result, to coordinate the supply chain and improve overall performance, this stream of literature often focuses on contracts[8], e.g., wholesale price contract ([30]), buy-back ([42]), and revenue-sharing contract ([12]). Many studies in this literature use the classical newsvendor model as the building block in a two-firm supply chain, and find remedies that can coordinate the system, while contracts are evaluated according to the following criteria: (1) coordination of supply chain[9], (2) arbitrary split of supply chain profits, and (3) administrative costs.

A significant body of research in this area has also studied the role of information sharing in achieving supply chain coordination. Studies in this domain investigated the value of sharing of information coming from the demand side, i.e., the value of giving the upstream members access to downstream information. For example, [14] analyzes the value of demand/inventory information sharing in a serial supply chain, while [11] study the value of inventory information from a downstream member's perspective in a one-warehouse, multi-retailer system. Our paper also focuses on information sharing from the downstream part of the supply chain, but considers a unique online advertising supply chain setup, which has not been previously analyzed in the extant

[8]Exhaustive reviews on this stream of literature can be found in [10] and [15].

[9]The buyer's decision should optimize the total supply chain profit.

literature.

We depart from the classical newsvendor setting, and study an online advertising supply chain with a different and a more complex structure involving vertical (intra-channel) as well as horizontal (inter-channel) interactions within the supply chain, and different pricing and transfer payment schemes. Due to the complexity of the online advertising supply chain and implementation issues, it is challenging to coordinate the supply chain. However, we show via simulation and numerical analysis that with information sharing and profit sharing schemes, the supply chain can increase its performance to a level that is comparable to that of the integrated case.

2.2.4 Empirical Research in Supply Chain Management

A growing body of literature has focused on empirical analyses of a gamut of issues in supply chains, including estimating unobservable cost parameters in the newsvendor model ([40]), the bullwhip effect ([5]), the value of information sharing ([18]), and the effect of vertical relationships on performance ([46]). [41] and [55] empirically verify the impact of supply chain risk propagation on production decisions and firm performance. Some of these issues also arise in our study, though in a digital setting, where vertical integration between the agency and the publishers is infeasible, and where not all players in the supply chain have equal visibility into various decisions. As a result, we develop solutions to these challenges that are distinct from those in the extant literature.

Another subset of supply chain management literature studies spillover effects or interdependencies among business practices, integration capabilities, and business resources (e.g., [32], [34], [37], [36], and [43]). [38], one of the earliest papers in this field, find that complexity in product design and vertical integration of production are complements, while [39] draw upon the theory of complementarity (see, for example, [35]), and examine complementarity among vertical integration decisions in automobile product development.

Our paper differs from prior supply chain management literature in three ways. First, our focus on synergies between online advertising channels differs from all of the aforementioned studies in supply chain management, which often focus on a supply chain of physical products such as automobiles. Our work on a unique digital supply chain not only provides a thorough analysis of the emerging performance-based advertising industry, but also opens up new avenues of research. Second, one of the reasons providing direct evidence of complementarities has been challenging in the extant literature is lack of information on costs and values of business practices, which ultimately hinders researchers in carrying out structural estimations. However, in our setting we are able to identify the exact costs, values, and other drivers of the decision maker's decisions as well as vertical and horizontal interactions within the digital supply chain, which enable us to empirically assess the presence of complementarities. Lastly, while most of the empirical research in supply chain management is descriptive in nature, we offer feasible recommendations to the agency which can increase campaign performance in terms of the agency's profit as well as that of the entire supply chain.

2.3 Decision Making Process

2.3.1 Decision Making Timeline

We first discuss the timeline of key decisions in the supply chain (Figure 2.2), which is based on our conversations with managers in the online advertising agency as well as observations from the data. In the campaign we study, the supply chain consists of the agency, a network in which the agency runs ads on a single publisher, and an exchange in which the agency buys impressions from a single broker. For notations that follow, we will use subscripts E and N for the exchange and the network respectively. In the beginning of a campaign, an advertiser and the agency negotiate the pay per action (PPA) for the exchange and the network, $p_{E,t}$ and $p_{N,t}$, the prices paid by the advertiser to the agency every time an advertisement leads to a specified measurable action (e.g., a sale, download, or subscription). These prices can vary over the course of a campaign through aperiodic renegotiation between the advertiser and the agency, and can also be different for conversions from the network and the exchange. Once the PPAs are set, the agency decides the following on a *daily* basis: (1) the number of impressions, $x_{E,t}$, it wishes to spend in the exchange given a (daily) CPM[10] and (2) the cost per action

[10]In reality, the agency specifies a daily spend (which is dosed throughout the day) and upper/lower bounds for bids for impressions in the demand side platform (DSP) which gives access to real-time-bidding of multiple sources of advertising inventory. Through experience with other campaigns, the agency has a good understanding of how many impressions its

(CPA) offered in the network, $w_{N,t}$. After observing the CPA offered by the agency, the publisher in the network decides on the number of impressions, $x_{N,t}$, to spend on its website. Finally, conversions, $y_{E,t}$ and $y_{N,t}$, are realized in both channels, the agency pays the CPA for each realized conversion to the network publisher, and the agency is paid by the advertiser for all realized conversions in the network and the exchange. It is critical to note that there is no exchange of information between the network and the exchange. We observed that cross-channel information flow is not allowed in the data management platforms used by the agency. Naturally, the publisher in the network is not privy to the transactions the agency makes in the exchange, and the broker in the exchange has no access to the negotiations taking place in the network between the publisher and the agency.

⟨⟨Figure 2.2 about here⟩⟩

The aforementioned decision making process in the supply chain allows us to formulate the interaction between the agency and a publisher in the network as a Stackelberg game. That is, the agency (the leader) first decides on the CPA, $w_{N,t}$, it is willing to pay to the publisher (as well as the number of impressions for the exchange, $x_{E,t}$), while the publisher in the network (the follower) decides on the number of impressions, $x_{N,t}$, only after observing the CPA, $w_{N,t}$. To characterize the subgame perfect Nash equilibrium (SPNE),

settings would result in given traffic and bidding information. As a simplifying assumption, we model the agency's daily spend decision as determining how many exchange impressions to spend given their belief of (average) CPM on a given day.

we first study the publisher's problem in the next subsection, and in Section 2.3.3 we study the agency's problem.

2.3.2 Publisher's Problem

We model the publisher's *view* of the conversion generation process by defining the number of conversions anticipated by the network, $y^P_{N,t}$, as a function of the number of impressions it spends in the network, $x_{N,t}$, as follows:

$$y^P_{N,t} = A^P_N x^{\alpha^P_N}_{N,t} e^{u^P_{N,t}} \tag{2.1}$$

where $u^P_{N,t}$ is a normally distributed random disturbance with mean 0 and variance $\left(\sigma^P_N\right)^2$, and superscript P refers to the publisher in the network. The uncertainty represents unpredictable variations in the number of conversions. The coefficient A^P_N is a positive constant, and α^P_N is the impression elasticity of conversion. We model the publisher's perspective of the conversion function as a single variable function of the number of impressions in the network, $x_{N,t}$, because the publisher has no visibility to the agency's decision on the exchange. The publisher's daily expected profit function maximization problem (PUB) is given as

$$\max_{x_{N,t}} \pi_N = w_{N,t}\, \mathbb{E}\left[y^P_{N,t} | x_{N,t}\right] - c_{N,t} x_{N,t} \tag{PUB}$$

where $w_{N,t}$ is the CPA paid by the agency at time t and $c_{N,t}$ is the opportunity cost of impressions per mille for the network publisher at time t. The publisher's opportunity cost is generally not known to the agency. To estimate the opportunity cost of the publisher, as we will show later in Section 2.4.1, we specify and estimate a simultaneous equations model of a profit maximizing publisher, which involves the publisher's conversion function and a profit maximizing condition, and estimate the opportunity cost as a function of total traffic in the network. Specifically, we define $c_{N,t} = k\left(x_{T,t}\right)^b$ where $x_{T,t}$ is the total traffic in the network, and k and b are constants to be estimated. As the traffic in the network increases, there are more impression opportunities for the publishers, thus ceteris paribus the opportunity cost of an impression decreases, i.e., $b < 0$. This is in sharp contrast to traditional advertising media such as radio or TV, where high traffic also leads to a high opportunity cost. Our data confirms the above intuition that the opportunity cost is a decreasing function of the total traffic on the network, i.e., the constant b is negative.

It can be shown that the publisher's optimal number of impressions, $x^*_{N,t}$, can be expressed as a function of the ratio of CPA and opportunity cost, ${w_{N,t}}/{c_{N,t}}$:

$$x^*_{N,t} = \left(A_N^P \alpha_N^P \left(\frac{w_{N,t}}{c_{N,t}} \right) e^{\frac{\left(\sigma_N^P\right)^2}{2}} \right)^{1/\left(1 - \alpha_N^P\right)} \tag{2.2}$$

Equation (2.2) implies the publisher uses more impressions in the campaign if the CPA is higher and/or the opportunity cost is lower.

2.3.3 Myopic Agency

We now investigate the agency's problem. The agency considers the best response of the publisher (i.e., how the publisher is expected to respond once the agency selects the CPA), and decides on the CPA in the network, and the number of impressions to use in the exchange that maximize its expected profit. Although managers in the agency have full visibility of decisions made in both the network and the exchange through the data management platform, we observed a decentralized structure of decision making and the complete lack of cross-channel communication and coordination within the agency, as a result of which the managers did not consider the interplay between the two channel structures in their decisions. Thus, while ideally the agency's perspective of the conversion processes in each channel should be functions of impressions in both the network and the exchange, to model the actual decision process we witnessed, we define the agency's *myopic* view of the conversion processes in each channel as

$$y_{E,t}^{M} = A_E^M x_{E,t}^{\alpha_E^M} e^{u_{E,t}^M},$$
$$y_{N,t}^{M} = A_N^M x_{N,t}^{\alpha_N^M} e^{u_{N,t}^M},$$

where $u_{E,t}^M$ and $u_{N,t}^M$ are normally distributed with zero means and variances $\left(\sigma_E^M\right)^2$ and $\left(\sigma_N^M\right)^2$ (and possibly nonzero covariances), and the superscript M refers to the myopic agency. As in Section 2.3.2, the coefficients A_E^M and A_N^M are positive constants, and α_E^M and α_N^M are the impression elasticities of conversion. The agency's expected profit maximization problem (AP) can be written as

$$\max_{x_{E,t}, w_{N,t}} \pi_A = p_{E,t}\, \mathbb{E}\left[y_{E,t}^M | x_{E,t}\right] - c_{E,t} x_{E,t} + \left(p_{N,t} - w_{N,t}\right) \mathbb{E}\left[y_{N,t}^M | x_{N,t}\right] \quad \text{(AP)}$$

where $p_{E,t}$ and $p_{N,t}$ are the price per action (PPA) in the exchange and the network respectively, $c_{E,t}$ is the cost per mille in the exchange.[11] Note also that $\mathbb{E}\left[y_{N,t}^M\right] = A_N^M \left(x_{N,t}^*\right)^{\alpha_N^M} e^{\frac{\left(\sigma_N^M\right)^2}{2}}$ where $x_{N,t}^*$ is decided by the network publisher as a function of the agency's CPA decision, $w_{N,t}$, as specified by Equation (2.2). The maximizing condition then yields the optimal number of impressions, $x_{E,t}^M$, which is a function of the ratio of PPA and CPM in the exchange:

$$x_{E,t}^M = \left(A_E^M \alpha_E^M \left(\frac{p_{E,t}}{c_{E,t}} \right) e^{\frac{\left(\sigma_E^M\right)^2}{2}} \right)^{1/\left(1 - \alpha_E^M\right)} \tag{2.3}$$

The agency spends more impressions in the exchange if the PPA is higher and/or the CPM is lower. The optimal CPA in the network, $w_{N,t}^M$, can be expressed as a constant times the PPA, $p_{N,t}$, in the network:

$$w_{N,t}^M = \frac{\alpha_N^M}{1 - \alpha_N^P + \alpha_N^M} \cdot p_{N,t} \tag{2.4}$$

We will use these insights from the myopic agency's problem, along with some adjustments the agency makes based on historical data, in choosing our instruments in the empirical analyses.

[11] We do not specify a budget for any of the models discussed in this paper. The agency is a well-funded publicly traded company, and it does not face any liquidity limitations forcing it to forgo profitable advertisement opportunities. As a matter of implementation, to control the campaigns, the agency sets a budget for each channel through their data management platform, and once the agency's spending reaches this budget limit, the ads will no longer be served. However, in reality these limits are increased if the ads served are profitable in terms of the resulting number of conversions.

2.4 Data and Empirical Model

We use data from a publicly traded multinational advertising agency that supplies advertisers and media partners with a wide range of online performance-based advertising solutions. Our study focuses on a single advertiser who runs multiple online campaigns on the advertising network and exchange through the agency. The advertiser provides an online service, and a conversion for the advertiser's campaign refers to a subscription to its website. Over the course of 382 days covered by the data, approximately 300 million impressions were used, and 45,202 conversions were realized, leading to a conversion rate of 0.015%. The advertising agency made a total revenue of $233,110 through this single campaign. Table 2.1 provides descriptive statistics for the campaign under consideration for 382 days.

⟨⟨Table 2.1 about here⟩⟩

Each row in our dataset represents a day in the campaign. A datum (row) consists of the number of impressions in the network and the exchange, the CPA bid to the network publisher, the number of conversions, revenue, cost, and profit from the publisher in the network and the broker in the exchange. Using these numbers, we calculate the PPA and conversion rate for the publisher in the network, and the PPA, CPM, and conversion rate for the broker in the exchange.[12] Moreover, we have the total number of impressions

[12]Since the actual CPM bids are not available in our analysis, we instead use the average daily CPM as a proxy for the actual bids.

the publisher in the network used for *all* the advertisers that are using its platform, and we use this number as a proxy for total traffic in the network. The ad creative used in the two channels are identical in content, size, type, position on the screen, and target geographic location.[13] If a customer sees the ad and clicks on it, he/she is directed to the advertiser's webpage where he/she is asked to subscribe to the website. The customer of course has a choice to leave the website, but if the customer provides his/her information to the website and creates an account his/her action is counted as a *conversion*.

The estimation process is carried out in three steps. First, we get estimates of the publisher's opportunity cost by specifying and estimating a simultaneous equations model of a profit maximizer, involving a demand equation and a performance equation. Second, we estimate the agency's decisions, the CPA for the network and the number of impressions for the exchange, using a set of instrumental variables based on the insights we gained from the previous section as well as conversations with the managers in the agency. Finally, using the estimated opportunity costs from the first step and the fitted values from the second step, we estimate the conversion functions for both channels. We describe each step in more detail below.

[13] Both channels use medium rectangle banner ads with the size of 300 x 250 placed on the right hand side of the screen *above the fold*, portion of a webpage that is visible to a customer when the page first loads. Also, ads on both channels are targeted to customers residing in the US.

2.4.1 Publisher's Decisions

The publisher's view of the conversion process is defined as Equation (2.1) and the publisher's expected profit function is given by (PUB) where, assuming that the disturbance in the conversion function is normally distributed with mean 0 and variance $\left(\sigma_N^P\right)^2$,

$$\mathbb{E}\left[y_{N,t}|x_{N,t}\right] = A_N^P x_{N,t}^{\alpha_N^P} e^{\frac{\left(\sigma_N^P\right)^2}{2}},$$

and $c_{N,t} = k\left(x_T\right)^b$. After a logarithmic transformation, the conversion process and the profit maximizing condition yield the following two equations for each day t:

$$\begin{aligned}
\ln y_{N,t} - \alpha_N^P \ln x_{N,t} &= \ln A_N^P + u_{N,t}^P, \\
\ln y_{N,t} - \ln x_{N,t} &= \phi_N' + \ln \frac{c_{N,t}}{w_{N,t}} + u_{N,t}^P + v_{N,t}
\end{aligned}$$

where $\phi_N' = \ln \frac{1}{\alpha_N^P} - \frac{\left(\sigma_N^P\right)^2}{2}$. The disturbance term $v_{N,t}$ is added in the second equation to explain deviations from the optimizing condition due to managerial errors. It is worth mentioning the difference in the nature of $u_{N,t}^P$ and $v_{N,t}$. The disturbance $u_{N,t}^P$ represents unpredictable variations in the conversion generation process, which is largely due to the variability in the customers' reaction to the ads. The disturbance term, $v_{N,t}$, represents the errors made by the managers in the agency. Hence, it is reasonable to assume that the correlation of those two error terms are zero. [60] note that the simple least squares estimators are consistent and unbiased under normality assumptions.

Rewriting the two equations leads to the following system of equations:

$$\ln y_{N,t} - \alpha_N^P \ln x_{N,t} - \ln A_N^P = u_{N,t}^P,$$
$$\left(\alpha_N^P - 1\right) \ln x_{N,t} - b \ln x_{T,t} + \ln w_{N,t} - \phi_N = v_{N,t}$$

where $\phi_N = \ln \frac{1}{A_N^P \alpha_N^P} - \frac{\left(\sigma_N^P\right)^2}{2} + \ln k$. We estimate the above system of equations through weighted least squares while imposing cross-equation restrictions on the coefficients. From the estimation results, we extract the publisher's decision rule, $x_{N,t}^*$ (see Equation (2.2)), the optimal number of impressions on day t, as a function of the network CPA, $w_{N,t}$, and the network opportunity cost, $c_{N,t}$, which again is a function total traffic in the network.

2.4.2 Optimal Agency's Decisions

We next investigate the agency's decisions, the CPA for the network, $w_{N,t}$, and the number of impressions in the exchange, $x_{E,t}$, on day t. Note that if the agency is myopic, then its decision rules are given as in Equations (2.3) and (2.4). We can verify this by estimating the following equations:

$$\ln x_{E,t} = \mu_0 + \mu_1 \ln \left(p_{E,t}/c_{E,t}\right) + \mu_2 LaggedConvsdiff_{E,t} + \mu_3 YearDummy + \epsilon_{E,t} \tag{2.5}$$

$$\ln w_{N,t} = \phi_0 + \phi_1 \ln p_{N,t} + \phi_2 LaggedCRdiff_{N,t} + \phi_3 YearDummy + \epsilon_{N,t} \tag{2.6}$$

There are some key features to the above equations that merit discussion. Based on our analysis in the previous sections, we assume that the exchange impressions depend on the ratio between the PPA and CPM in the exchange while the CPA in the network depends on the PPA in the network. Third, we acknowledge that the PPA for both channels and the CPM for the exchange may not be exogenously determined. We find, however, that the PPA for both channels are not directly related to any of the variables we constructed from the data set. Nevertheless, we observe that there are two distinct jumps over the course of the campaign and minimal variations elsewhere. Hence, we introduce two dummy variables, Shock 1 and Shock 2, to represent three different levels of the PPA. As for the exchange CPM, we observe that it largely depends on previous period's CPM, $c_{E,t-1}$, and the conversion rate in the exchange, $CR_{E,t-1}$. The higher the conversion rate in the previous period, the CPM bid by the agency becomes larger in order to secure advertising slots. Lastly, the disturbance terms, $\epsilon_{E,t}$ and $\epsilon_{N,t}$, refer to the agency's error in its attempt to satisfy the profit maximizing conditions. To estimate the agency's decisions we specify an unrestricted variance-covariance matrix for the error terms and estimate the system of equations via a three-stage least squares estimation (3SLS) procedure. From this estimation step, we calculate the fitted values of the (log of) network CPA, $\widehat{\ln w_{N,t}}$, and the (log of) exchange impressions, $\widehat{\ln x_{E,t}}$, and use them in our final estimation step discussed in the next subsection.

2.4.3 Conversion Functions

The last step of the estimation requires estimating the conversions from each channel on day t, $y_{E,t}$ and $y_{N,t}$. After log transformation the conversion functions can be written as follows:

$$\ln y_{E,t} = \ln A_E + \alpha_E \widehat{\ln x_{E,t}} + \beta_E \ln x_{N,t} + u_{E,t} \tag{2.7}$$

$$\ln y_{N,t} = \ln A_N + \alpha_N \ln x_{N,t} + \beta_N \widehat{\ln x_{E,t}} + u_{N,t} \tag{2.8}$$

If the sum of the coefficients in each equation is less than 1, as our empirical analysis will later show, the conversion process exhibits decreasing returns to scale; in this case each incremental amount of online advertising causes a lesser increase in conversions, which may be explained as a result of advertising saturation. The coefficients that are of particular interest to us are β_N and β_E. These coefficient reflect the inter-channel or horizontal spillover effects, so if the coefficients turn out to be positive, then we can infer that there are positive spillover effects from one channel to the other. The error terms, $u_{E,t}$ and $u_{N,t}$, capture the randomness in consumers' reactions. Note that we use the fitted values for the (log of) exchange impressions, $\widehat{\ln x_{E,t}}$, which we derived from the previous step of estimation. Our analysis in Section 2.3.2 indicates that the number of impressions on the publisher's website is a function of the proportion of the CPA, $w_{N,t}$, and the opportunity cost, $c_{N,t}$, in the network (see Equation (2.2)). We also observe that the publisher's decision depends on historically how many conversions it has delivered in the past. For this reason, we also include the lagged number of conversions in the network, $y_{N,t-1}$, as one of the instruments for the network impressions, $x_{N,t}$. Again, to estimate the conversion functions we specify an unrestricted variance-covariance matrix for the error terms and estimate the system of equations via 3SLS. Figure 2.3 visualizes our three-step estimation procedures.

⟨⟨Figure 2.3 about here⟩⟩

2.4.4 Exclusion Restrictions and Identification

Since the endogenous variables are correlated with the disturbances, the OLS estimators are inconsistent. As emphasized in [3] and [4], we can use an instrumental variables (IV) framework to disentangle complementarity from clustered organizational practices. This approach can provide a consistent estimate of the synergy coefficients if one can find measures correlated with the endogenous variables but uncorrelated with the error term. Our analysis in the previous section yields a natural set of instruments: factors driving the agency's decisions, which are the network CPA, network opportunity cost, exchange PPA, and the exchange CPM. Moreover, we consider additional covariates that are extracted from direct observation from the data and our discussion with the agency. Table 2.2 below summarizes the endogenous variables included in estimation equations and the excluded exogenous variables in each of the equations.

⟨⟨Table 2.2 about here⟩⟩

The justification for exclusion restrictions are discussed in Sections 2.4.2 and 2.4.3. Since the number of endogenous variables is strictly less than the number of excluded exogenous variables, the order conditions are satisfied. Moreover, the systems of equations trivially satisfy the rank conditions and thus we are able to identify all the coefficients in the simultaneous equations models.

2.5 Empirical Results

The estimation results for the publisher's problem is presented in Table 2.3.

⟨⟨Table 2.3 about here⟩⟩

From the parameter estimates from the publisher's problem, we can extract the publisher's opportunity cost as an inverse function of the total number of impressions (which is used as a proxy for total traffic on the publisher's website). Summary statistics for the opportunity cost along with the exchange CPM is provided in Table 2.4.

⟨⟨Table 2.4 about here⟩⟩

Table 2.5 presents the estimation results for the agency's decisions via the 3SLS method. The positive and significant coefficients for $\ln\left({}^{p_{E,t}}/{}_{c_{E,t}}\right)$ and $\ln p_{N,t}$ indicate that the number of impressions for the exchange and the CPA for the network are positively affected by the proportion of PPA and CPM in the exchange and the PPA in the network, respectively. Moreover, if the number of conversions from the exchange is higher than that of the network, the exchange manager exerts more effort in the exchange, and if the conversion rate in the network is higher than that in the exchange, the network manager increases the CPA for the network. Table 2.6 provides the estimation results for the conversion functions. We find that the cross-channel spillovers are indeed positive and significant, providing strong evidence of complementarities between the two channels. A one-unit change in the log of network (exchange) impressions induces a 10% (7%) increase in the log of exchange (network) conversions. Moreover, the coefficients satisfy the condition for supermodularity of the agency's expected profit function, i.e., the profit increase that results from increasing both the number of impressions in the exchange and the CPA in the network is *greater* than the sum of profit increases that result from increasing either one in isolation. This shows that decisions for both channels should be made in tandem after incorporating their interdependencies in the optimization model, and that a failure to do so may result in suboptimal decision-making, i.e., possibly underspending on some actions.

Our results regarding the complementarity of conversions in the impressions from the two channels adds to the nascent body of literature on complementarities in physical supply chains (e.g., [38], [39]) that have found synergies between practices, integration initiatives and contracting.

$\langle\langle$Table 2.5 and 2.6 about here$\rangle\rangle$

2.6 Supply Chain Coordination

In this section, we devise several schemes that can potentially boost the profit level of the agency as well as the supply chain compared to the baseline case, the status quo of the agency's decision making discussed in Section 2.3. We extract 7 days of data from the original dataset, and use the PPA in each channel, CPM in the exchange, and the opportunity cost in the network to generate decisions, conversions, and expected profit to compare the performances of each scheme.

2.6.1 Benchmark and Baseline

We first describe the ideal setting where the advertising network and the agency are integrated. It is important to note that this benchmark is not practical, since the agency does not own either the large publisher in the network (or those in the real-time-bidding exchange). This analysis only serves as a reference point to compare the performances of the actual and proposed configurations of the supply chain. Moreover, due to the horizontal and vertical complexities present in the supply chain, it is challenging, if not infeasible, to achieve the same level of profit as the benchmark case with any other arrangement; however, along with considerations of potential interactions between channels, in Sections 2.6.3 through 2.6.5 we devise schemes that will considerably boost the supply chain's profit over the baseline case.

In the benchmark scenario, the agency decides the number of impressions for both channels, $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$. We use notations $\bar{y}_{E,t}$ and $\bar{y}_{N,t}$ for the number of conversions in each channel, where the over-bar refers to the integrated supply chain. If the agency were to consider the inter-channel interactions in its decisions, its view of the conversion generation processes in each channel will ideally include impressions from both channels, in which case we can model the conversion processes as

$$\bar{y}_{E,t} = \bar{A}_E \bar{x}_{E,t}^{\bar{\alpha}_E} x_{N,t}^{\bar{\beta}_E} e^{\bar{u}_{E,t}}$$
$$\bar{y}_{N,t} = \bar{A}_N \bar{x}_{N,t}^{\bar{\alpha}_N} x_{E,t}^{\bar{\beta}_N} e^{\bar{u}_{N,t}}$$

where $\bar{u}_{E,t}$ and $\bar{u}_{N,t}$ are normally distributed disturbances with 0 mean, variances $\bar{\sigma}_E^2$ and $\bar{\sigma}_N^2$, and possibly nonzero covariance. The agency's expected profit maximization problem for the integrated case, which we refer to as INT, can be written as

$$\max_{\bar{x}_{E,t},\bar{x}_{N,t}} \bar{\pi}_A = p_{E,t}\,\mathbb{E}\left[\bar{y}_{E,t}|\bar{x}_{E,t},\bar{x}_{N,t}\right] - c_{E,t}\bar{x}_{E,t} + p_{N,t}\,\mathbb{E}\left[\bar{y}_{N,t}|\bar{x}_{N,t},\bar{x}_{E,t}\right] - c_{N,t}\bar{x}_{N,t}. \tag{INT}$$

We derive the following result for the integrated chain case:

Proposition 1. *Assume that $\bar{\beta}_E$ and $\bar{\beta}_N$ are non-negative, i.e., there are non-negative spillovers across channels. Then the following statements hold.*

1. *The expected profit function π_A is supermodular in $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$.*
2. *The optimal decisions, $\bar{x}_{E,t}^*$ and $\bar{x}_{N,t}^*$, can be selected such that they are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$ and $c_{N,t}$.*

The first statement in the above Proposition shows that supermodularity at the conversion level leads directly to the supermodularity of the expected profit function. This is due to the fact that the marginal cost of increasing $\bar{x}_{E,t}$ is independent of $\bar{x}_{N,t}$ and vice versa (see Appendix for details), which means the cross-partial derivative of $\bar{\pi}_A$ with respect to $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$ is always non-negative. We will show in later sections (Sections 2.6.2 and 2.6.3) that this result does not necessarily extend to alternative supply chain arrangements. The second statement asserts that there exist agency's optimal decisions that are non-decreasing in the other channel's PPA, and are non-increasing in the other channel's cost of impressions. While this result may seem intuitive, we show in sections 2.6.2 and 2.6.3 that this may not necessarily be true in other supply chain arrangements. That is, the Stackelberg structure which we introduce in the next section may induce the agency to be more aggressive and select the network CPA, $w_{N,t}$, to be larger than the network PPA, $p_{N,t}$, in which case the agency's actions may be increasing in the network opportunity cost, $c_{N,t}$. Moreover, if the solution pair is unique, it is guaranteed that the properties in the second statement hold[14]. However, if the spillover coefficients, $\bar{\beta}_E$ and $\bar{\beta}_N$, are non-positive, the following result can be derived:

Corollary 1.1. *Assume that $\bar{\beta}_E$ and $\bar{\beta}_N$ are non-positive, i.e., there are non-positive spillovers across channels. Then the following statements hold.*

1. *The expected profit function π_A is submodular in $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$.*
2. *The optimal decision, $\bar{x}^*_{E,t}$ $\left(\bar{x}^*_{N,t}\right)$, can be selected such that it is non-decreasing (non-increasing) in $p_{E,t}$ and $c_{N,t}$, and is non-increasing (non-*

[14] Topkis's Monotonicity Theorem ([53], Theorem 2.8.3) suggests that if the action set is a compact rectangle in a Euclidean space, the parameter set is a partially ordered set, and the profit function is supermodular in the actions, and has increasing differences in the actions and the parameters, then the solution set is a non-empty compact sublattice for all parameters, and that the greatest and least elements in the solution set is non-decreasing in the parameters. Moreover, if the solution is unique, then the solution is a non-decreasing function.

decreasing) in $p_{N,t}$ and $c_{E,t}$.

This result indicates that submodularity on the conversion level is carried over to the expected profit function, and that the agency's optimal decisions are decreasing in the other channel's PPA. Note also that even if one of the coefficients, $\bar{\beta}_E$ and $\bar{\beta}_N$, is zero, the results in Proposition 1 and Corollary 1.1 still hold. If one of the coefficients is negative (positive), then the supermodularity (submodularity) of the expected profit function in Proposition 1 (Corollary 1.1) is not always guaranteed, as it depends on the relative magnitude of the PPAs and the conversion function parameters. In the next section, we discuss how circumstances can change in a vertically segmented supply chain with an independent publisher in the network.

To simulate outputs in the integrated supply chain scenario, we begin by computing the agency's optimal decisions. Specifically, in this setting the agency decides the number of impressions for both channels, $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$. The agency's decisions are based on the first order conditions of INT, where the number of conversions are generated from Equations (2.7) and (2.8). The profit levels of the agency, the network, and the supply chain are calculated based on these decisions and the number of conversions from both channels. Next, we compare it with the supply chain profit generated by the agency's decisions in the baseline case, i.e., the myopic agency's case. This profit is obtained through simulation, and it is carried out in the following steps. First, in order to extract the agency's view of the impression elasticities, we follow the steps in Sections 2.4.1, 2.4.2, and 2.4.3, but in estimating the conversion functions we exclude the impression variable from the other channel. That is, we estimate the following two conversion equations:

$$\ln y_{E,t} = \ln A_E^M + \alpha_E^M \widehat{\ln x_{E,t}} + u_{E,t}^M$$
$$\ln y_{N,t} = \ln A_N^M + \alpha_N^M \ln x_{N,t} + u_{N,t}^M$$

This approach extracts the agency's beliefs about the impression elasticities, α_E and α_N. Table 2.7 summarizes the results. Based on these coefficient estimates, we generate the agency's decisions, the exchange impressions, $x_{E,t}$, and the network CPA, $w_{N,t}$, using Equations (2.3) and (2.4). The publisher's decision, the number of impressions in the network, $x_{N,t}$, is then generated based on Equation (2.2). The realized profit levels of the agency, the network, and the supply chain are calculated accordingly.

The decisions from the agency and the network as well as the profit levels for the agency and the supply chain are provided in Table 2.8.[15] We observe that the optimal profit levels for the agency and the supply chain are more than four times the profits in the baseline scenario. The operating level in each channel has also increased dramatically relative to the benchmark case. We acknowledge these are out-of-range figures, and that such operating levels may not be achieved in reality due to various constraints; however, these numbers suggest that there is a significant upside potential, and that it

[15]We also provide decisions and profit levels from the actual data set in Table 8. The difference between the actual data and the baseline case numbers are due to managerial errors within the agency and the network publisher.

is of interest to study how the agency can improve its performance through acknowledging complementarity and utilizing different contract schemes.

⟨⟨Table 2.7 and 2.8 about here⟩⟩

2.6.2 Informed Agency Model

As noted earlier, multiple studies (e.g., [1], [31], [57], etc.)[16] provide theoretical and empirical support for interactions across channels. If the agency were to consider the inter-channel interactions in its decisions (we will call this agency the *informed* agency), its view of the conversion process in a channel will include impressions from the other channel as in the integrated supply chain case:

$$y_{E,t} = A_E x_{E,t}^{\alpha_E} x_{N,t}^{\beta_E} e^{u_{E,t}} \tag{2.9}$$

$$y_{N,t} = A_N x_{N,t}^{\alpha_N} x_{E,t}^{\beta_N} e^{u_{N,t}} \tag{2.10}$$

where $u_{E,t}$ and $u_{N,t}$ are normally distributed with zero means, variances σ_E^2 and σ_N^2, and nonzero covariances. The informed agency's expected profit maximization problem (IA) can be written as

$$\max_{x_{E,t}, w_{N,t}} \pi_A = p_{E,t} \, \mathbb{E} \left[y_{E,t} | x_{E,t}, x_{N,t} \right] - c_{E,t} x_{E,t} + (p_{N,t} - w_{N,t}) \, \mathbb{E} \left[y_{N,t} | x_{N,t}, x_{E,t} \right] \tag{IA}$$

[16] These studies analyze individual consumer level data, and study complementarities between online advertising formats such as display and search. Their focus is not on the interactions within the advertising supply chain.

where $\mathbb{E}\left[y_{E,t}\right] = A_E x_{E,t}^{\alpha_E}\left(x_{N,t}^*\right)^{\beta_E} e^{\frac{\sigma_E^2}{2}}$, $\mathbb{E}\left[y_{N,t}\right] = A_N\left(x_{N,t}^*\right)^{\alpha_N} x_{E,t}^{\beta_N} e^{\frac{\sigma_N^2}{2}}$ and $x_{N,t}^*$ is decided by the network publisher as a function of the agency's CPA decision, $w_{N,t}$, as specified by Equation (2.2). Note that the agency's expected profit is based on Equations (2.9) and (2.10) while using the optimal decision made by the network publisher from Equation (2.2), and optimizing over this expected profit function determines $w_{N,t}^*$ and $x_{E,t}^*$. Thus maximizing the expected profit with these conversion processes involves finding solutions to a nonlinear system of equations given below:

$$\frac{\partial \pi_A}{\partial x_{E,t}} = p_{E,t}\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}} - c_{E,t} + \left(p_{N,t} - w_{N,t}\right)\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} = 0 \tag{2.11}$$

$$\frac{\partial \pi_A}{\partial x_{N,t}} = p_{E,t}\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}}\frac{\partial x_{N,t}^*}{\partial w_{N,t}} + \left(p_{N,t} - w_{N,t}\right)\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x_{N,t}^*}{\partial w_{N,t}} - \mathbb{E}\left[y_{N,t}\right] = 0 \tag{2.12}$$

Although solving for the solutions requires numerical procedures, we can show that under certain conditions, the expected profit function is supermodular in the agency's decisions, the CPA in the network and the number of impressions in the exchange. We find that even if the conversions in the network and exchange are supermodular in the impressions, i.e., the output elasticities, β_E and β_N, are non-negative, the supermodularity of the agency's expected profit function with respect to the agency's decisions, $x_{E,t}$ and $w_{N,t}$, is not guaranteed.

Proposition 2. *Assume that β_E and β_N are non-negative, i.e., there are non-negative spillovers across channels, and that*

$$p_{E,t}\frac{\partial^2\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x_{N,t}^*}{\partial w_{N,t}} + \left(p_{N,t} - w_{N,t}\right)\frac{\partial^2\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x_{N,t}^*}{\partial w_{N,t}} - \frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} \geq 0. \tag{A1}$$

Then the expected profit function π_A is supermodular in $x_{E,t}$ and $w_{N,t}$, and the optimal decisions, $x^*_{E,t}$ and $w^*_{N,t}$, can be selected such that they are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$. Further, if $w_{N,t} \leq p_{N,t}$, then $x^*_{E,t}$ is non-increasing in $c_{N,t}$.

The above result suggests that if the spillover effects are non-negative, and the sum of marginal benefits from both channels of increasing an action while the other action is also increasing dominates the spillover effect of the exchange impressions to the network (i.e., the third term in (A1)), then the agency's decisions are complements, and the optimal decisions, $x^*_{E,t}$ and $w^*_{N,t}$, are non-decreasing in the PPA in the network and the exchange, and are non-increasing in the CPM in the exchange, $c_{E,t}$. Moreover, if the CPA in the network is less than or equal to the PPA in the network, then $x^*_{E,t}$ is non-increasing in the opportunity cost in the network, $c_{N,t}$. Note that in Assumption (A1), all the partials and cross partials are non-negative if the spillover effects, β_E and β_N, are non-negative:

$$\underbrace{p_{E,t}\frac{\partial^2\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}}_{(+)}+\left(p_{N,t}-w_{N,t}\right)\underbrace{\frac{\partial^2\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}}_{(+)}-\underbrace{\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}}}_{(+)}$$

However, the equation can turn negative, hence turning the expected profit function submodular with respect to the agency's decisions, depending on the relative magnitude of the components. For example, if the spillover effect of the exchange impressions to the expected number of conversions in the network, $\partial\mathbb{E}[y_{N,t}]/\partial x_{E,t}$, is too large, holding all others fixed, then it may become more profitable for the agency to reduce its effort in the exchange while increasing its effort in the network (and vice versa). The same conclusion holds if the efficiencies of the network impressions, α_N and α_N^P, are too low, holding all others fixed. This result is in contrast with that in the benchmark case, and can be attributed to the vertical interactions in this model. In the benchmark case, the marginal cost of increasing the network impressions is precisely the opportunity cost of impressions, $c_{N,t}$, and this cost remains the same regardless of the increase in exchange impressions for the campaign, i.e., the marginal cost of increasing the campaign impressions on the network while increasing the exchange impressions is zero. However, in the informed agency case the marginal cost of increasing the network CPA depends indirectly on the network CPA itself through the expected value of network conversions, $\mathbb{E}[y_{N,t}]$,(see Equation (2.12)). Moreover, the marginal cost of increasing the network CPA while increasing the exchange impressions depend both on the network CPA and the exchange impressions. Therefore, the marginal cost of increasing both actions is no longer zero. We can also derive the following result for the case where the spillover effects, β_E and β_N, are non-positive.

Corollary 2.1. *Assume that β_E and β_N are non-positive, i.e., there are non-positive spillovers across channels, and that*

$$p_{E,t}\frac{\partial^2\,\mathbb{E}[y_{E,t}]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+(p_{N,t}-w_{N,t})\,\frac{\partial^2\,\mathbb{E}[y_{N,t}]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}-\frac{\partial\,\mathbb{E}[y_{N,t}]}{\partial x_{E,t}}\leq 0.$$

Then the expected profit function π_A is submodular in $x_{E,t}$ and $w_{N,t}$, and the optimal decisions, $x^*_{E,t}\left(w^*_{N,t}\right)$, can be selected such that it is non-decreasing in

$p_{E,t}$ $(p_{N,t})$, and is non-increasing in $p_{N,t}$ $(p_{E,t})$ and $c_{E,t}$. Further, if $w_{N,t} \leq p_{N,t}$, then $x^*_{E,t}$ is non-decreasing in $c_{N,t}$.

The above corollary suggests that the agency's expected profit function can turn supermodular even if the spillover effects, β_E and β_N, are negative, depending on the magnitude of ${}^{\partial \mathbb{E}[y_{N,t}]}/_{\partial x_{E,t}}$, holding all others fixed. We also find that if the coefficient β_N is zero, the supermodularity (submodularity) of the expected profit function is guaranteed for non-negative (non-positive) β_E. However, if the coefficient β_E is zero, or if one of the coefficients, β_E and β_N, is negative (positive) while the other is non-negative (non-positive), then the supermodularity (submodularity) of the expected profit function is not always guaranteed.

We now describe the simulation procedure for the informed agency scenario. Using the impression elasticities from Section 2.4 for the conversions, we generate the agency's decisions, $x_{E,t}$ and $w_{N,t}$, through Equations (2.11) and (2.12). These decisions represent the agency's decisions had it acknowledged complementarities between channels and incorporated them into their optimization. Plugging these decisions back into the conversion functions (2.7) and (2.8), we get the number of conversions from both channels, and we can calculate the profit levels for the agency, the network, and the supply chain. We can see from Table 2.8 that in the informed agency case, the agency's profit level increases by 30% over the baseline, and the supply chain's profit level almost doubles. That is, the agency can exploit its informational advantage in better allocating its budget to both channels not only to achieve higher profit but also to increase that of the other players in the supply chain. Acknowledging complementarities between channels clearly boosts the profit levels of the agency and the supply chain, but there exists a large gap between the profit levels of the informed agency case and the benchmark. To help reduce this difference, we devise several implementable supply chain contracts involving information and profit sharing.

2.6.3 Information Sharing

We first consider an information sharing scheme for the agency. As mentioned before, the publisher in the network does not have visibility into the transactions between the exchange and the agency. Thus, the publisher cannot assess potential interactions with the exchange. Our analysis shows that if the agency chooses to share the transaction information on the exchange with the publisher, the supply chain profit can more than triple. Specifically, with the updated information from the agency, the publisher's perspective on the conversion process now becomes

$$\ln y_{N,t} = \ln A_N + \alpha_N \ln x_{N,t} + \beta_N \ln x_{E,t} + u_{N,t}$$

where the coefficients, α_N and β_N, are from Equation (2.8). The publisher's optimal decision now becomes a function of both the number of impressions in the exchange and the ratio of CPA and opportunity cost as given below:

$$x^*_{N,t} = \left(A_N \alpha_N \left(\frac{w_{N,t}}{c_{N,t}} \right) x_{E,t}^{\beta_N} e^{\frac{(\sigma_N)^2}{2}} \right)^{1/(1-\alpha_N)} \tag{2.13}$$

Moreover, since the sign of the coefficient, β_N, is positive, i.e., there is a positive spillover effect from the exchange to the network, the publisher's optimal decision, $x^*_{N,t}(x_{E,t})$, is increasing in the number of impressions the agency chooses in the exchange. Given the publisher's decision, the agency's expected profit maximization problem is changed accordingly. We can show that under certain conditions, the agency's expected profit function is supermodular in the agency's decisions. We will use *IS* to represent the information sharing scenario.

Proposition 3. *Assume that β_E and β_N are non-negative, i.e., there are non-negative spillovers across channels, and that*

$$\begin{aligned} p_{E,t} &\left(\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t} \partial x_{N,t}} \frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x^2_{N,t}} \frac{\partial x^*_{N,t}}{\partial x_{E,t}} \frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial \mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}} \frac{\partial^2 x^*_{N,t}}{\partial x_{E,t} \partial w_{N,t}} \right) \\ &- \left(\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} + \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}} \frac{\partial x^*_{N,t}}{\partial x_{E,t}} \right) + (p_{N,t} - w_{N,t}) \\ &\left(\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t} \partial x_{N,t}} \frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x^2_{N,t}} \frac{\partial x^*_{N,t}}{\partial x_{E,t}} \frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}} \frac{\partial^2 x^*_{N,t}}{\partial x_{E,t} \partial w_{N,t}} \right) \geq 0. \end{aligned} \tag{A2}$$

Then the expected profit function π_A is supermodular in $x_{E,t}$ and $w_{N,t}$, and the optimal decisions, $x^{IS}_{E,t}$ and $w^{IS}_{N,t}$, can be selected such that they are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$. Further, if $w_{N,t} \leq p_{N,t}$, then $x^{IS}_{E,t}$ is non-increasing in $c_{N,t}$.

The condition A2 appears more complex than condition A1 of Proposition 2, which is largely due to the fact that in the information sharing case the publisher's decision, $x^*_{N,t}$, also depends on the exchange impressions, $x_{E,t}$.

However, the intuition behind the result remains similar. That is, if the cross-channel spillover effects are non-negative, and the sum of the marginal benefits from both channels of increasing an action while the other action is also increasing dominates the spillover effect of the exchange impressions to the network, then the agency's decisions are complements, and the optimal decisions, $x_{E,t}^{IS}$ and $w_{N,t}^{IS}$, are non-decreasing in the PPA in both channels, and are non-increasing in the CPM in the exchange. It is worth noting, however, that the spillover effect of the exchange impressions to the network can be larger than that of the informed agency case due to the extra term, $\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}} \frac{\partial x_{N,t}^{*}}{\partial x_{E,t}}$. The exchange impressions can affect the network conversions through the publisher's impressions since in this case the publisher's impressions depend on the exchange impressions. More precisely, the marginal cost of increasing both actions simultaneously can be larger than that of the informed agency case because the change in the marginal cost of increasing the network CPA as the agency increases the exchange impressions is affected not only by the *direct* impact of exchange impressions on the expected number of network conversions but also by the *indirect* impact of exchange impressions on the expected number of network conversions through network impressions (see proof in the appendix for detailed expressions). The above result also suggests that if the CPA in the network is less than or equal to the PPA in the network, then the optimal number of exchange impressions is non-increasing in the opportunity cost in the network. We find that the agency's profit can be increased by 166% over the baseline case through information sharing (see Table 2.8). We can also derive the following result for the case where the spillover effects, β_E and β_N, are non-positive.

Corollary 3.1. *Assume that β_E and β_N are non-positive, i.e., there are non-positive spillovers across channels, and that*

$$p_{E,t}\left(\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+\frac{\partial \mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right)$$
$$-\left(\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}}+\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right)+\left(p_{N,t}-w_{N,t}\right)$$
$$\left(\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right)\leq 0.$$

Then the expected profit function π_A is submodular in $x_{E,t}$ and $w_{N,t}$, and the optimal decisions, $x^{IS}_{E,t}$ $\left(w^{IS}_{N,t}\right)$, can be selected such that it is non-decreasing in $p_{E,t}$ $(p_{N,t})$, and is non-increasing in $p_{N,t}$ $(p_{E,t})$ and $c_{E,t}$. Further, if $w_{N,t} \leq p_{N,t}$, then $x^{IS}_{E,t}$ is non-increasing in $c_{N,t}$.

The intuition behind the above result is similar to that of Corollary 2.1.

2.6.4 Publisher's Profit Sharing Contract

The supply chain profit can be increased even more by devising profit sharing schemes. First, suppose that a contract is negotiated in which the publisher shares a portion, $(1-\gamma)$, of its profit with the agency in exchange for the information shared by the agency. The publisher's expected profit maximization problem is then given as

$$\max_{x_{N,t}} \pi_N = \gamma\left(w_{N,t}\,\mathbb{E}\left[y_{N,t}|x_{N,t},x_{E,t}\right]-c_{N,t}x_{N,t}\right)$$

while the agency's expected profit maximization problem (PPS) is given as

$$\max_{x_{E,t}, w_{N,t}} \pi_A = p_{E,t}\,\mathbb{E}\left[y_{E,t}\right] - c_{E,t}x_{E,t} + \left(p_{N,t} - w_{N,t}\right)\mathbb{E}\left[y_{N,t}\right] \\ + \left(1-\gamma\right)\left(w_{N,t}\,\mathbb{E}\left[y_{N,t}\right] - c_{N,t}x_{N,t}\right) \quad \text{(PPS)}$$

We emphasize that the expected number of conversions on the network as estimated by the network, $\mathbb{E}\left[y_{N,t}\right]$, now includes knowledge about the impressions in the agency bought in the exchange. Notice that if the publisher's share γ is 1, then the agency's problem reduces to the informed agency's problem, and if γ is 0, the problem reduces to the integrated chain scenario. In this profit sharing scheme, who determines the share, γ, can be critical. Suppose the agency has higher bargaining power than the publisher, i.e., the agency chooses γ.[17] It is straightforward to see that the publisher's decision, $x^*_{N,t}$, does not *directly* depend on γ (but it depends indirectly on γ through the network CPA, $w_{N,t}$), since given the updated information about the exchange impressions the publisher's decision is given as in Equation (2.13). Note, however, that the agency's decisions do depend on the choice of γ, i.e., the exchange impressions, $x_{E,t}$, and the network CPA, $w_{N,t}$, are functions of γ. We can derive the following proposition.

Proposition 4. *Let $k = \frac{c_{N,t}x^*_{N,t}}{w_{N,t}\,\mathbb{E}\left[y_{N,t}\right]}$, the publisher's cost-benefit ratio (cost over expected revenue). If $k \leq 1$, then the optimal decisions, x_E^{PPS} and w_N^{PPS}, can be selected such that it is non-decreasing in γ.*

[17] For example, if the publisher is a local grocery store website, the agency will have a stronger bargaining power.

The above proposition suggests that if the publisher's cost-benefit ratio is less than or equal to 1, i.e., if the cost of impressions is less than or equal to the expected revenue, then the agency's optimal decisions are non-decreasing in the agency's share of the publisher's profit.

Now suppose the publisher has more bargaining power than the agency, i.e., the publisher chooses its share γ.[18] Then it becomes a three-stage game between the agency and the publisher. The publisher announces γ, and the agency decides on the exchange impressions and the network CPA in response. Finally, based on the agency's decisions the publisher determines the network impressions. Table 2.8 presents the simulation results for both scenarios. We find that the agency's profit can be increased by 166% to 272% over the baseline case, depending on the bargaining power between the agency and the publisher. We can also verify that the higher the agency's share, $(1 - \gamma)$, the higher the operating levels of the agency. Our discussion is presented in the form of a party, the one with greater bargaining power, making the decision about γ and imposing on the other; in general the value of γ is determined by a bargaining process.

2.6.5 Agency's Profit Sharing Contract

The second profit sharing scheme is related to altering the agency's payment structure to the publisher and the agency's decision variables. Specif-

[18] An example of this would be the agency working with large publishers such as Facebook or AOL.

ically, the agency pays the publisher on a CPM basis instead of the CPA method, and offers to share with the publisher the revenue from the advertiser less the cost of impressions. The agency's expected profit maximization problem, which we refer to as APS, can be written as

$$\max_{x_{E,t}} \pi_A = p_{E,t}\,\mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right] - c_{E,t}x_{E,t} + \gamma\left(p_{N,t}\,\mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right] - c_{N,t}x_{N,t}\right) \tag{APS}$$

while the publisher's problem is now changed to

$$\max_{x_{N,t}} \pi_N = (1-\gamma)\left(p_{N,t}\,\mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right] - c_{N,t}x_{N,t}\right)$$

Notice that in this scenario the agency has one decision variable, the exchange impressions $x_{E,t}$. If γ is equal to 1, the agency's problem reduces to the integrated chain case. As in the publisher's profit sharing case, the bargaining power between the two parties plays a major role. Suppose the agency has more bargaining power than the publisher, i.e., the agency chooses γ. Then again, the publisher's decision, $x^*_{N,t}$, does not directly depend on γ while the agency's decision, $x_{E,t}$, does. We can derive the following proposition.

Proposition 5. *Let $k' = \frac{c_{N,t}x^*_{N,t}}{p_{N,t}\,\mathbb{E}\left[y_{N,t}\right]}$, the cost-benefit ratio (cost over expected revenue). If $k' \leq 1$, then the optimal decision, x_E^{APS}, can be selected such that it is non-decreasing in γ.*

This result implies that if cost of impressions is less than or equal to 1, then the agency's optimal decision is non-decreasing in the agency's share of the network profit. If the publisher has more bargaining power than the agency, then the interaction between the agency and the publisher again becomes a three-stage game. We find that the agency's profit can be increased by 163% - 239% depending on the preset profit sharing scheme, γ. Table 2.8 summarizes the results. We can also verify the result in Proposition 5 by observing that the agency's decision, $x_{E,t}^{APS}$, is higher for higher γ.

2.6.6 Discussion

As seen in Section 2.4, there exist positive spillover effects across channels. However, as we established from, for example, Proposition 2, positive spillover effects do not necessarily guarantee economic complementarity, i.e., supermodularity of the agency's expected profit function in its decisions, $x_{E,t}$ and $w_{N,t}$. For such complementarity to be achieved, certain conditions have to be met for each scenario. We verified that the conditions A1 and A2 were met for each respective scenario in our simulation analysis. However, it was not always the case that the optimal network CPA is less than the network PPA. In the informed agency case, the optimal network CPA was *always* higher than the network PPA (notice in Table 2.8 that the mean network CPA is the highest in the informed agency case), and in all other cases including the information sharing scenario and the publisher's profit sharing scenario, the network CPA was higher than the network PPA at least once. This implies that in order to exploit the effect of complementarities, it may sometimes be profitable for the agency to forgo profitability in one channel and extract profit from the other.

Table 2.8 also includes in the last column the agency's share of the supply chain profit in each of the scenarios discussed earlier. In the baseline scenario, the agency keeps 89% of the supply chain profit while the publisher in the network gets 11%. In the informed agency case, however, the agency's share drops down to 65%. While the agency's profit shows an increase of 30%, the network publisher gets to keep a larger portion of the supply chain profit increase mainly because the agency is giving up profit in the network by setting the network CPA higher than the network PPA in order to fully exploit the channel complementarities. Once the agency starts sharing information with the publisher, the publisher now incorporates channel complementarities in its decision making so that the agency becomes less aggressive than in the informed case. Consequently, the agency's share of the profit increases to 69%. With publisher's or agency's profit sharing in addition to information sharing, the agency can extract a larger profit share as much as 84%, depending on the relative magnitude of the bargaining powers between the agency and the publisher.

While the simulation results show that there is potential of increasing profit levels of the players in the supply chain, we acknowledge that the optimal operating levels may not be implementable in practice due to various reasons. In Table 2.9 we provide how the profit levels of the agency, the network publisher, and the supply chain will change in response to incremental increases (2% - 20%) in the agency's decisions. The first three columns show the percentage increase in profits when the agency increases its decisions *si-*

multaneously, and the latter three columns represent percentage increase in the *sum* of profits when the agency increases its decisions individually. It is evident that percentage increase in the former case is larger than that of the latter, signifying the presence of economic complementarity.

⟨⟨Table 2.9 about here⟩⟩

Finally, Table 2.10 summarizes the agency's share of the supply chain profit[19] in cases with different parameters than the original ones. Each row in the table represents cases with different sets of parameters while each column corresponds to the baseline scenario, the informed agency scenario, and the information sharing scenario. The symmetric case is where we assume that the network has the same impression elasticities of conversions, i.e., $\alpha_E^M = \alpha_N^M = 0.6890$, $\alpha_E = \alpha_N = 0.6910$, and $\beta_E = \beta_N = 0.1016$, and the same costs, i.e., $c_{E,t} = c_{N,t}$, as the exchange. The case with different spillovers refers to the scenario where the coefficients β_E and β_N are different (β_N is back to its original value, 0.0731). We find that in this case the agency's share increases in all three scenarios. Although due to the drop in spillover from the network to the exchange the profit levels decrease for both the agency and the network publisher, the publisher sees a relatively higher drop in the profit than does the agency because there is less spillover effect from the exchange to the network than in the symmetric case. In the case with different efficiencies

[19]Unlike the numbers in Table 2.8 which are based on simulated conversions and profit functions, the numbers in Table 2.10 are based on the *expected* values of the counterparts.

(α_N and α_N^M are back to their original values) the agency's share decreases in all three scenarios compared to the symmetric case. In this case, the increase in the efficiencies favorably affects the profit levels for both the agency and the network publisher, but the agency enjoys a relatively lower increase in the profit levels than does the network publisher. When we let the cost of impressions in the network back to its original values (which are higher than the exchange CPM), the agency's profit shares increase in all three scenarios compared to the symmetric case. As in the case with different spillovers, this is because the publisher sees a relatively higher drop in the profit than does the agency. The network publisher's profit critically depends on the cost of impressions in the network while the agency's profit function does not directly depend on it. Note that the status quo (or the original case) is the combination of all three variants considered. Based on the numbers in original case, we can infer that the effect of different efficiencies is dominated by the effect of different spillovers and costs.

⟨⟨Table 2.10 about here⟩⟩

2.7 Conclusion

In spite of a substantial body of knowledge in online advertising on interactions between channels and consumers, there is no analysis of the online advertising supply chain, which has essentially been treated as a black box in the extant literature. We analyzed the vertical interactions within a channel as well as horizontal synergies between channel structures in an online advertising supply chain. Based on the insights from our discussions with the managers in the agency, we modeled and validated the presence of strong complementarities through simultaneous equations estimation.

Our study provides the rationale for a new economic role of the agency in terms of information sharing and coordination activities. We showed that information and profit sharing can substantially boost the performance of the supply chain; yet the IT systems used by various players in the supply chain are not integrated, and therefore do not provide visibility across channels. For example, the publisher has no knowledge of choices made in the real-time bidding systems. The agency, being in a position to observe both channels, can provide the necessary information to the publisher to boost its own as well as the supply chain's performance.

Our results also underscore the need to redesign the agency's current organizational structure from specializing in media buying (the status quo) to campaign based management, where one (or more) decision maker(s) in charge of a campaign is (are) made aware of horizontal synergies as well as informational aspects involved in the supply chain. Figure 2.4 illustrates the agency's current and our proposed organization structures.

⟨⟨Figure 2.4 about here⟩⟩

The status quo has separate and specialized groups of managers for the exchange and the network. The current organizational structure reflects a rationalization of the role of the agency in the supply chain based on economies of scale in media buying and lowering search and transaction costs for the advertisers; it is well-suited for exploring new publishers and brokers in the two channels, and having access to their online content. However, our results suggest a broader rationalization of the economic role of the agency, one that is based on the utilization of information, and one which provides a significant increase in profit potential by considering cross-channel interactions when making decisions for each channel. Hence, we propose that the agency group together managers in charge of similar campaigns, and either integrate the network and the exchange management teams or ensure through interventions that they understand the impact of synergies on their decision making and agency profits. Such structural changes will involve additional costs of cross-training managers as well as business process changes to move from a silo mode of operation to a more integrated focus on the overall supply chain; nevertheless, our numerical results also show that there is a significant upside potential from such an initiative. Our findings also show that setting the right incentive system is also key to improved performance. The current incentive system in the agency is based on the number of realized conversions each manager extracts from his/her campaign. Naturally, the channel managers tend to put less emphasis on the cost of impressions and even less on the potential interactions across channels. A gainsharing system for each campaign may motivate the managers to better work as a team and improve their campaign performances.

Tables and Figures

	Total	Exchange Total	Network Total	Exchange Daily Average	Network Daily Average
Mille Impressions	297,611	250,244	47,366	655	124
Conversions	45,202	28,713	16,489	75	43
Revenue	233,110	149,027	84,083	390	220
Cost	146,547	89,244	57,303	234	150
ROI (%)	59	67	48		
Average PPA				5.14	5.14
Average CPA					3.42
Average CPM				0.73	
Conversion Rate (%)	0.0152	0.0115	0.0348		

Table 2.1: Descriptive Statistics

	Endogenous Variables	Exogenous Variables
(2.9)	$\ln p_{E,t}$, $\ln c_{E,t}$	Shock 1, Shock 2, $\ln c_{E,t-1}$, $CR_{E,t-1}$
(2.10)	$\ln p_{N,t}$	Shock 1, Shock 2
(2.11,2.12)	$\ln x_{N,t}$	$\ln (\widehat{w_{N,t}}/c_{N,t})$, $y_{N,t-1}$, $Year\ Dummy$

Table 2.2: Endogenous and Exogenous Variables

Variables	Conversion Function	FOC
Intercept	0.9831	-5.3485***
$\ln x_{N,t}$	0.7428***	-
$\ln w_{N,t}$	-	3.8880***
$\ln x_{T,t}$	-	0.8911***
Adjusted R^2	0.8624	0.4511

$^{\cdot}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 2.3: Coefficient Estimates for Publisher's Problem

	Min	Median	Mean	Max
Network's Opportunity Cost	0.6620	1.0922	1.2704	5.1881
Exchange CPM	0.0500	0.6200	0.7272	3.3100

Table 2.4: Network's Opportunity Cost and Exchange CPM

Variables	$\ln x_{E,t}$	$\ln w_{N,t}$
Intercept	0.6647^{*}	-1.1207^{***}
$\ln \left(p_{E,t}/c_{E,t}\right)$	1.4605^{***}	-
LaggedConvsdiff	2.2062^{***}	-
$\ln p_{N,t}$	-	1.4416^{***}
LaggedCRdiff	-	0.0125^{*}
Year Dummy	-1.8996^{***}	-0.0517^{***}
Adjusted R^2	0.6775	0.4950

$^{\cdot}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 2.5: Coefficient Estimates for Agency's Decisions

Variables	$\ln y_{E,t}$	$\ln y_{N,t}$
Intercept	0.5946^{*}	0.5543^{***}
$\widehat{\ln x_{E,t}}$	0.6910^{***}	0.0731^{***}
$\ln x_{N,t}$	0.1016^{**}	0.8115^{***}
Adjusted R^2	0.5267	0.8647

$^{\cdot}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 2.6: Coefficient Estimates on Conversion Functions

Variables	$\ln y_{E,t}$	$\ln y_{N,t}$
Intercept	0.8695	0.7819^{***}
$\widehat{\ln x_{E,t}}$	0.6890^{***}	-
$\ln x_{N,t}$	-	0.8048^{***}
Adjusted R^2	0.5239	0.8539

$^{\cdot}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 2.7: Coefficient Estimates on Conversion Functions for Baseline Scenario

	γ	$x_{E,t}$	$w_{N,t}$	$x_{N,t}$	Agency Profit (%)	Overall Profit (%)	Agency Share (%)
Data		1,482	3.02	198	-	-	97
Baseline		1,912	3.82	180	100	100	89
IA		4,122	5.38	696	129.65	177.65	65
IS		9,109	4.84	4,188	266.18	342.96	69
PPS	0.999*	9,112	4.84	4,192	266.30	343.05	69
	0.322**	11,754	5.34	7,731	372.11	393.87	84
APS	0.001*	9,268	-	5,055	262.67	360.40	65
	0.531**	10,475	-	5,306	339.36	364.17	83
INT		16,650	-	10,912	453.90	403.51	100

*Publisher's best option; **Agency's best option

Table 2.8: Simulation Results

	Simultaneous Increase			Separate Increase		
	Agency	Network	Overall	Agency	Network	Overall
2%	2.19	10.08	3.03	1.93	9.56	2.74
4%	4.16	21.16	5.98	4.09	19.02	5.69
6%	6.11	31.62	8.84	5.74	29.99	8.33
8%	8.20	44.76	12.12	7.75	42.44	11.46
10%	10.09	56.71	15.08	9.37	56.53	14.43
12%	11.93	71.42	18.31	11.13	70.01	17.44
14%	13.74	85.49	21.43	12.46	83.94	20.12
16%	15.40	101.87	24.67	13.72	99.64	22.92
18%	17.04	118.78	27.94	15.13	115.69	25.91
20%	18.57	136.88	31.24	16.18	133.67	28.77

Table 2.9: Percentage Increases in Profit Levels for Increases in Agency's Decisions

	Baseline	Informed Agency	Information Sharing
Symmetric	74.96%	56.66%	66.70%
Different Spillovers	82.33%	65.28%	71.79%
Different Efficiencies	68.31%	51.06%	60.75%
Different Costs	82.28%	60.35%	70.20%
Original	89.29%	66.77%	71.04%

Table 2.10: Agency's Share of Supply Chain Profit

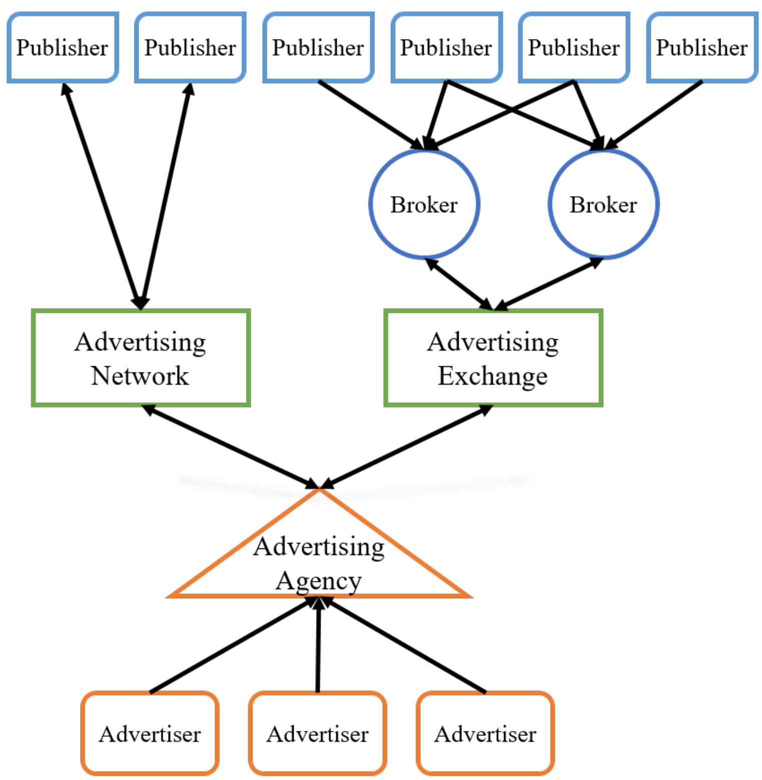

Figure 2.1: An Online Advertising Supply Chain

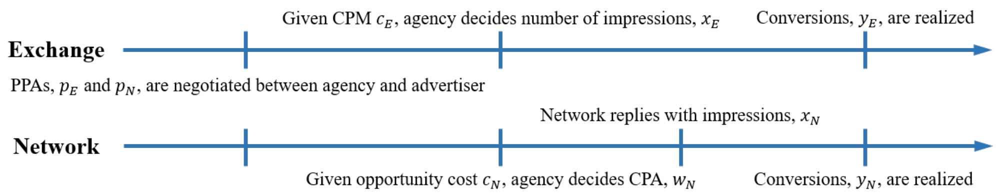

Figure 2.2: Agency's Decision Making Timeline

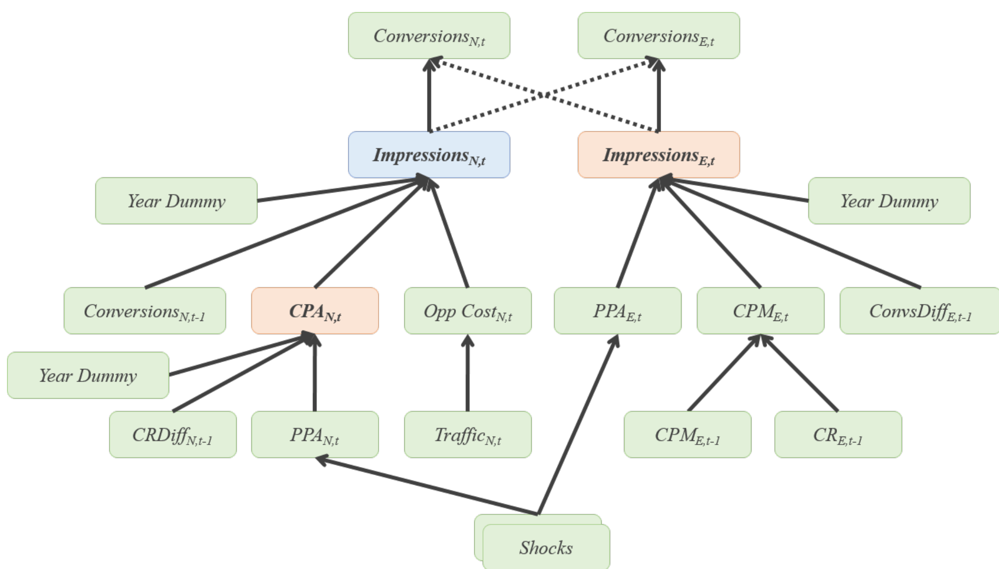

Figure 2.3: Empirical Model

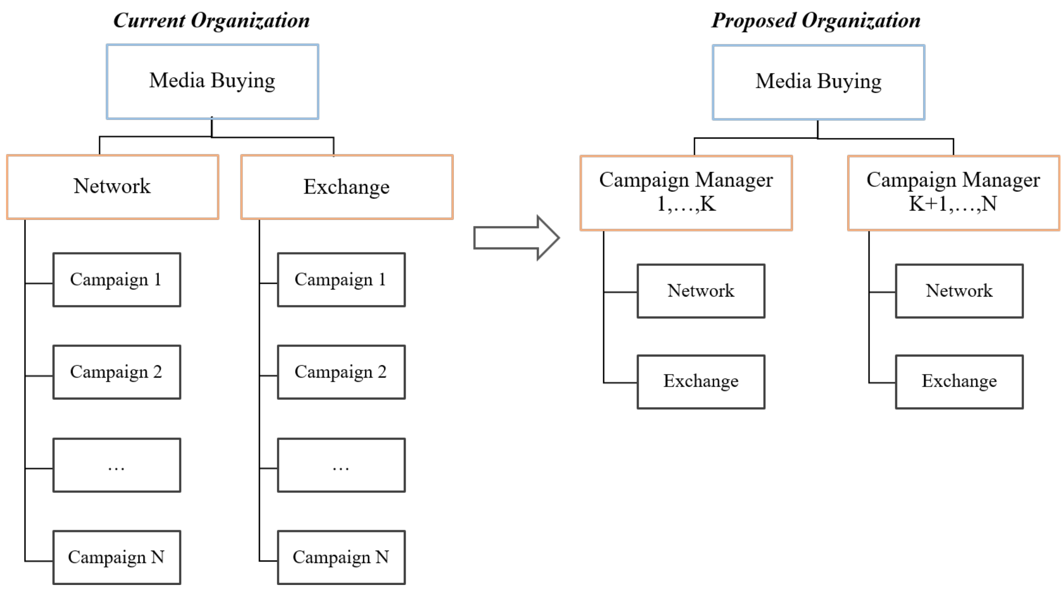

Figure 2.4: Reorganization of the Agency

Chapter 3

The Effect of Organizational Design on Exploiting Complementarities

3.1 Introduction

The management literature in organizational design has investigated how firms should be organized (centralized, decentralized, or hybrid) to successfully pursue business strategies in different contexts. The information systems literature, while replete with studies on testing complementarities among IT practices, has focused less on how organizational design or governance mode affects a firm's ability to exploit the synergy effects (with very few exceptions such as [19] and [52]). In this study, we empirically investigate how centralized versus decentralized decision making structure in a firm can affect its ability to exploit synergies among business practices and how this lead to various performance measures in the context of online advertising.

We study an online advertising *supply chain*, which involves an online advertising agency who attempts to sell a product using two different channels; a network of large publishers, and an exchange with real-time bidding for impressions on publishers' websites. The advertising network is a closed group of agencies and publishers akin to a privately traded market, where prices are determined through individual negotiations, while the advertising exchange is a technology-driven platform that facilitates buying and selling of impressions inventory with prices determined by a real-time-bidding (RTB) algorithm. Major ad exchange operators include Google DoubleClick Ad Exchange, Microsoft Ad Exchange, and OpenX. The agency buys impressions from the two channels, and a conversion is realized when a customer carries out a measurable action, e.g., a customer makes a purchase, provides information, or downloads an app, in response to an impression. Then the agency is paid a price for each realized conversion from the advertiser, and the agency's profit margin is the difference between the revenue they receive from the advertiser and the cost they pay for impressions spent on both channels. Conversion generation processes in the two channels may not be independent as the impressions are targeted at the same set of customers. However, it is not clear if there are synergies present or if there is a substitution effect. That is, we do not know if increasing impressions in one channel will improve or hinder the effectiveness of the impressions in the other channel.

When campaign managers launch a campaign, each manager is initially in charge of one advertising channel with its own platform, and as managers assimilate their respective platforms they tend to cross-train themselves and manage multiple channels simultaneously. Hence, our data from a single campaign naturally consists of two main parts: one that is produced by two managers each controlling one of the two channels, and the other that is generated by a single manager who is in charge of making decisions in two different channels. More specifically, the ad campaign we study is initially run by a single manager (Manager A); that is, Manager A is in charge of making decisions in both the network and the exchange channels. After some time, another manager (Manager B) becomes involved in the campaign, and s/he controls the decisions in the network channel, while Manager A keeps control of the exchange. Then in the last phase of the campaign, Manager B takes control of both channels. The data thus provides a natural experimental setting through which we can econometrically verify which decision mode (centralized or decentralized) creates super-additive value synergies from cross-channel complementarities. We initially focus on quantifying complementarities between impressions on the network of publishers and exchange platform in terms of their impact on conversions in the two channels using a simultaneous equations estimation approach involving instrumental variables. We then explore the effect of centralization versus decentralization on recognizing and exploiting the synergy effects by simulating the agency's decisions assuming different decision making structures and belief systems of the agency.

We find that the centralized mode, with a single manager for both channels, better recognizes and incorporates complementarities between the two channels (the network and the exchange) into its decision making, while the decentralized mode with two managers, one in charge of each channel, fails to recognize the effect of complementarities and systematically fails to optimally allocate the advertising efforts, leading to less profit for the agency. Our analysis demonstrates that the decision structure can make a difference in the performance due to more effective exploitation of cross-channel synergies in an online advertising campaign.

3.2 Literature Review

To the best of our knowledge, this paper is the first to rigorously quantify the impact of organizational design on firm performance in the presence of complementarities between business practices. The relevant literature to our work can be divided into two main areas: organizational design and synergies in online advertising.

3.2.1 Literature on Organizational Design

Many scholarly articles in the management field have addressed the issue of firms' organizational design in relation to various firm activities (e.g., [49], [7], [47], [48], [56]). [56] study the impact of top management teams' integrative complexity and decentralization of decision making on corporate social performance, while [47] analyze how three different organizational structures (centralized, decentralized, and temporarily decentralized) moderate the balance of exploration and exploitation of a firm. With more focus on the information systems function, [7] propose theory of predicting a centralized, decentralized, or compromise design solution between the corporate and business-unit levels of management for systems development. We add to this literature by, building on the economic theory of complementarities, quantitatively verifying how organizational design is associated with exploiting the effect of complementarities between two different business decisions. Another factor that distinguishes our model is the unique context of online advertising and the dataset that resembles an experimental setting, which allow us to empirically verify the change in performance measures associated with the change in organizational structure.

3.2.2 Synergies in Online Advertising

Our study largely draws upon the nascent literature of synergies among multiple modes or formats of online advertising. For example, [58] show that organic and sponsored search advertising have positive interdependence while [1] estimate a hidden Markov model of consumer behavior and find that display and search ads affect customers differently based on their states in the purchasing process. [57] use a mutually exciting point process to show that display advertisements stimulate subsequent visits to other advertising formats and eventual conversions. [25] show that combining web and mobile display ads performs better than when either web or mobile is used in isolation. [29] find that display ads increase search clicks and conversions.

Our work differs from the above literature in two distinct ways. First, prior literature does not consider the decision making structure and/or process by the managers who control ad campaigns. Instead, these studies primarily focus on the interactions between ads on websites and individuals visiting them, and analyze, for example, strategies to increase the likelihood of clicks or conversions. By contrast, our study analyzes the decision making process of the campaign managers and how the firm's organizational design affects various performance metrics of the online ad campaign we study. Second, most of the above-mentioned literature primarily focuses on quantifying the synergy effects whereas our paper tackles the problem of how in the presence of complentarities a firm can benefit from exploiting the effect by possibly reorganizing its decision making structure. Thus, our work complements the extant literature and provides detailed analysis to gain a better understanding the association between organizational design and recognition/exploitation of the effect of complementarities.

3.3 The Online Advertising Supply Chain

In the online advertising campaign we study, the supply chain consists of an ad agency, a network in which the agency runs ads on a single publisher, and a real-time-bidding exchange in which the agency buys impressions from a single broker who aggregates advertising slots from various publishers and sell them to advertising agencies through the exchange platform. In the beginning of an ad campaign, the advertiser and the agency negotiate the pay per action (PPA)[1] for the exchange and the network, $p_{E,t}$ and $p_{N,t}$, the prices paid by the advertiser to the agency per each conversion. These prices can vary over the course of a campaign through renegotiation between the advertiser and the agency, and can also be different for conversions from the network and

[1]PPA is the amount of money that the agency receives from an advertiser every time an advertisement leads to a specified action (e.g., a sale, click, download, or subscription).

the exchange, though such variation appears to be minimal in our data. Once the PPAs are set, the agency decides the number of impressions, $x_{E,t}$ and $x_{N,t}$, it wishes to spend in the exchange and the network, respectively, given the cost per mille (CPM)[2] in each channel. Then, conversions, $y_{E,t}$ and $y_{N,t}$, are realized in both channels, and the agency is paid by the advertiser for all realized conversions in the network and the exchange.

3.3.1 Informed Agency

Multiple studies (e.g., [1], [31], [57], etc.) provide theoretical and empirical support for interactions across channels. The prima facie justification of positive interactions or complementarities is that repeated exposure across multiple websites may entice a consumer to click on an ad, and even buy the advertiser's product, as studied in the literature on interaction between different formats. If the agency considers the inter-channel interactions in its decisions, its view of the conversion process in a channel will include impressions from the other channel as

$$y_{E,t} = A_E x_{E,t}^{\alpha_E} x_{N,t}^{\beta_E} e^{u_E} \tag{3.1}$$

$$y_{N,t} = A_N x_{N,t}^{\alpha_N} x_{E,t}^{\beta_N} e^{u_N} \tag{3.2}$$

where u_E and u_N are normally distributed with zero means, variances σ_N^2 and σ_E^2, and possibly nonzero covariances. The agency's expected profit maximiza-

[2]The underlying assumption for the exchange is that every participant in the exchange is a price taker. Moreover, the agency has a fair assessment of the CPM through extensive learning before the launching of a campaign.

tion problem can be written as

$$\max_{x_{E,t},x_{N,t}} \pi_A = p_{E,t}\,\mathbb{E}\left[y_{E,t}|x_{E,t},x_{N,t}\right] - c_{E,t}x_{E,t} + p_{N,t}\,\mathbb{E}\left[y_{N,t}|x_{N,t},x_{E,t}\right] - c_{N,t}x_{N,t}.$$

where $\mathbb{E}\left[y_{E,t}\right] = A_E x_{E,t}^{\alpha_E} x_{N,t}^{\beta_E} e^{\left(\sigma_E^2/2\right)}$, $\mathbb{E}\left[y_{N,t}\right] = A_N x_{N,t}^{\alpha_N} x_{E,t}^{\beta_N} e^{\left(\sigma_N^2/2\right)}$. Maximizing the expected profit with these conversion processes involves finding solutions to a nonlinear system of equations given below:

$$\frac{\partial \pi_A}{\partial x_{E,t}} = p_{E,t}\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}} - c_{E,t} + p_{N,t}\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} = 0 \tag{3.3}$$

$$\frac{\partial \pi_A}{\partial x_{N,t}} = p_{E,t}\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}} + p_{N,t}\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}} - c_{N,t} = 0 \tag{3.4}$$

3.3.2 Myopic Agency

Although managers in the agency technically have visibility of decisions made in both the network and the exchange through their data management platform, we observed a decentralized structure of decision making and the lack of cross-channel communication and coordination within the agency especially during the second phase of the campaign where two managers are at play, as a result of which the managers did not consider the interaction between the two channels in their decisions. Thus, while ideally the agency's perspective of the conversion processes in each channel should be functions of impressions in both the network and the exchange (as shown in Equations (3.1) and (3.2), to model the decision process we witnessed, we define the agency's *naive* view of the conversion processes in each channel as

$$y_{E,t} = A_E^M x_{E,t}^{\alpha_E^M} e^{u_E^M} \tag{3.5}$$

$$y_{N,t} = A_N^M x_{N,t}^{\alpha_N^M} e^{u_N^M} \tag{3.6}$$

where u_N^M and u_E^M are normally distributed with zero means and variances $\left(\sigma_N^M\right)^2$ and $\left(\sigma_E^M\right)^2$ (and possibly nonzero covariances), and where the superscript M represents the naive agency. The maximizing conditions then yield the optimal numbers of impressions, $x_{E,t}^M$ and $x_{N,t}^M$, which are functions of the price cost ratios in their respective channels:

$$x_{E,t}^M = \left(A_E^M \alpha_E^M \left(\frac{p_{E,t}}{c_{E,t}} \right) e^{\frac{\left(\sigma_E^M\right)^2}{2}} \right)^{\frac{1}{\left(1-\alpha_E^M\right)}} \tag{3.7}$$

$$x_{N,t}^M = \left(A_N^M \alpha_N^M \left(\frac{p_{N,t}}{c_{N,t}} \right) e^{\frac{\left(\sigma_N^M\right)^2}{2}} \right)^{\frac{1}{\left(1-\alpha_N^M\right)}} \tag{3.8}$$

3.4 The Data and Estimation

3.4.1 The Data

We use data from a publicly traded, multinational online advertising agency. We focus on an advertiser who ran an online campaign using a publisher on the advertising network and a broker in the advertising exchange through the agency for 180 days. The advertiser provides an online service, and a conversion for the advertiser's campaign refers to a subscription to its website. Manager A was in charge of both channels for the first 57 days of the campaign, and from the 58$^{\text{th}}$ day to the 120$^{\text{th}}$ day, Manager B took control over the network, while Manager A continued to be in charge of the exchange. Lastly, from the 121st day Manager B took control over both channels until the end of the campaign. Table 3.1 provides descriptive statistics for the campaign under consideration.

⟨⟨Table 3.1 about here⟩⟩

Each observation in our dataset consists of a day of operations. A datum (row) consists of the number of impressions in the network and the exchange, the PPA values, the number of conversions, revenue, cost, and profit from the publisher in the network and the broker in the exchange. We then use this information to construct additional variables: daily conversion rates (CR), CPMs, and effective costs per action (eCPA) for each channel. The typical industry definition of CR is the number of conversions from an ad divided by the number of clicks on the ad, i.e., CR = conversions/clicks. In this paper, however, we will define CR to be the number of conversions divided by the number of impressions. Also, since the actual CPM bids are not available in our data, we instead use the average daily CPM as a proxy for the actual bids. eCPA is defined as the cost of impressions divided by the number of *realized* conversions resulting from the impressions. The ad creative used in the two channels are identical in terms of the content, size, type, position on the screen, and target geographic location. Both channels used medium rectangle banner ads with the size of 300 x 250 placed on the right hand side of the screen above the fold, portion of a webpage that is visible to a customer when the page first loads. Ads on both channels were targeted to customers residing in the US. If a customer sees the ad and clicks on it, he/she is directed to the advertiser's webpage where he/she is asked to subscribe to the website. The customer of course has a choice to leave the website, but if the customer provides his/her information to the website and creates an account his/her action is counted as a conversion.

3.4.2 The Empirical Framework

We use the classical Cobb-Douglas specification to test complementarities between the exchange and the network channels, which has been the most commonly used model in research relating inputs to output and testing complementarities. The estimation process is carried out in two main steps. In the first step, we estimate dynamic regression models (AR(1)) for the CPMs in each channel with respective lagged conversion rates as regressors and use the fitted values from the models to use as bases for the managers' beliefs about the CPMs in making decisions for the impressions on any given day. Second, we estimate the conversion functions using an iterative three-stage least squares (3SLS, [61]). [6], among others, suggests that iterative generalized least squares approaches lead to global maximum likelihood estimates in the limit, which implies that full information methods such as 3SLS can be used in place of full information maximum likelihood (FIML). We describe each step in more detail below.

3.4.2.1 Dynamic Regression Model for CPM

When making decisions on the number of impressions to spend in each channel, the managers in the agency form a belief about the daily average CPM on a given day with the information at hand. To mimic this belief forming process, we estimate dynamic regression models for the CPMs and use the fitted values from the models as the agency managers' belief. More specifically, we observe that the CPM values are generally sticky, i.e., they largely depend on previous period's CPM, $c_{E,t-1}$ and $c_{N,t-1}$. They also depend on previous period's conversion rates, $CR_{E,t-1}$ and $CR_{N,t-1}$, respectively. The underlying logic here is that the higher the conversion rate in the previous period, the larger the CPM bid by the agency in order to secure advertising slots.

We thus estimate a first-order autoregressive (AR(1)) dynamic regression model for each channel with previous period's conversion rate as a regressor. For the exchange we have

$$
\begin{aligned}
c_{E,t} &= \theta_0 + \theta_1 CR_{E,t-1} + \eta_t, \\
\eta_t &= \phi \eta_{t-1} + \epsilon_t
\end{aligned}
$$

where ϕ is the first-order autocorrelation coefficient and ϵ_t is white noise which represents the part of CPM that cannot be explained by this model. We can write a similar equation for the network CPM. The fitted values from these models, $\widehat{c_{E,t}}$ and $\widehat{c_{N,t}}$, represent the agency's belief about the CPMs which are used to make decisions on the impressions.

3.4.2.2 Estimation of Conversion Functions

Next, we estimate the conversions from each channel on day t, $y_{E,t}$ and $y_{N,t}$ via 3SLS. After log transformation the conversion functions can be written as follows:

$$\ln y_{E,t} = \ln A_E + \alpha_E \ln x_{E,t} + \beta_E \ln x_{N,t} + \gamma_E Phase2 + \delta_E Phase3 + \kappa_E Trend + u_{E,t} \tag{3.9}$$

$$\ln y_{N,t} = \ln A_N + \alpha_N \ln x_{N,t} + \beta_N \ln x_{E,t} + \gamma_N Phase2 + \delta_N Phase3 + \kappa_N Trend + u_{N,t} \tag{3.10}$$

If the sum of the elasticities, α and β, in each equation is less than one, as our empirical analysis will later show, the conversion process exhibits decreasing returns to scale; in this case each incremental amount of online advertising causes a lesser increase in conversions, which may be explained as a result of advertising saturation. The coefficients, β_E and β_N, reflect the inter-channel or horizontal spillover effects, so if the coefficients turn out to be positive, then we can infer that there are positive spillover effects from one channel to the other. We also include dummy variables, $Phase2$ and $Phase3$, which are 1 in the corresponding phases and 0 otherwise. Finally we include the trend variable to control for trend in our regressions.

It is reasonable to think that in each channel's conversion equation its own impression variable is potentially endogenous. That is, exchange (network) impressions, $x_{E,t}$ ($x_{N,t}$), are correlated with the disturbance, $u_{E,t}$ ($u_{N,t}$), because it is very likely that there are variables that affect both the conversions and impressions in the same channel simultaneously. For example, increase in total traffic on a website may drive both the impressions and conversions to increase. Impression variable from the other channel has less of this concern because it is quite unlikely that there exists a common variable that simultaneously affects, for example, both the exchange side conversions and the network publisher's impressions. Due to the above-mentioned endogeneity issue, the OLS estimators are inconsistent. As emphasized in [3] and [4], we can use an instrumental variables (IV) framework to disentangle complementarity from organizational practices. This approach can provide a consistent estimate of the synergy coefficients if one can find instrumental variables that satisfy relevance and exclusion restriction assumptions ([24]), i.e., they should be correlated with the corresponding endogenous regressor (relevance) and uncorrelated with the error term (exclusion restriction).

Our analyses in Section 3.3.2 yield a natural set of instrumental variables. In particular, the number of impressions in each channel is a function of the respective price-cost ratios in each channel (see Equations (3.7) and (3.8)). Therefore we use these ratios, $p_{E,t}/\widehat{c_{E,t}}$ and $p_{N,t}/\widehat{c_{N,t}}$, as instruments for the impression variables, $x_{E,t}$ and $x_{N,t}$, respectively. Both of these ratios naturally satisfy the relevance condition.[3] In addition, we would expect that these ratios affect the conversions only through the impressions because the general audience of the ads have no information about these ratios when they are making their decisions to *convert*. It is also important to note that the error terms,

[3]There of course exists the assumption that the agency is a profit maximizer.

$u_{E,t}$ and $u_{N,t}$, may be correlated, i.e., the conversions from the exchange and network can be simultaneously affected by a common unobserved exogenous factors such as changes in consumers' reactions due to an unknown event (the correlation between exchange conversions and network conversions is about 0.53). To allow the error terms to be contemporaneously correlated, we use the three-stage least squares (3SLS, [61]) approach, a feasible generalized least squares (FGLS) version of the two-stage least squares (2SLS, [2]) estimation, which leads to more efficient estimates. More specifically, we use the following procedure:

Stage 1: Estimate via ordinary least squares (OLS) the endogenous independent variables, $x_{E,t}$ in Equation (3.9) and $x_{N,t}$ in Equation (3.10), using instrumental variables, i.e., $p_{E,t}/\widehat{c_{E,t}}$ for $x_{E,t}$ and $p_{N,t}/\widehat{c_{N,t}}$ for $x_{N,t}$, as well as the dummy variables and trend variable in Equations (3.9) and (3.10). We then calculate the predicted endogenous independent variables, $\widehat{\ln x_{E,t}}$ and $\widehat{\ln x_{N,t}}$.

$$\ln x_{E,t} = \rho_E + \lambda_E \ln(p_{E,t}/\widehat{c_{E,t}}) + \ln x_{N,t} + Controls + \nu_{E,t}$$
$$\ln x_{N,t} = \rho_N + \lambda_N \ln(p_{N,t}/\widehat{c_{N,t}}) + \ln x_{E,t} + Controls + \nu_{N,t}$$

Stage 2: Using the predicted endogenous variables from Stage 1, we estimate the coefficients in Equations (3.9) and (3.10) via OLS.

$$\ln y_{E,t} = \ln A_E + \alpha_E \widehat{\ln x_{E,t}} + \beta_E \ln x_{N,t} + Controls + u_{E,t}$$
$$\ln y_{N,t} = \ln A_N + \alpha_N \widehat{\ln x_{N,t}} + \beta_N \ln x_{E,t} + Controls + u_{N,t}$$

We use these 2SLS estimates to predict residuals in the system of equations estimation, which are then used to compute the contemporaneous residual covariance matrix.

Stage 3: Compute the general least squares estimators of the system of equations. We iterate the above process till the coefficient estimates have converged.

In Section 3.4.3.1, we provide additional analysis to show the validity of the instruments.

3.4.3 The Estimation Results

Table 3.2 summarizes the key results from each of the 3SLS estimations, which indicate that there are significant synergy effects between the network and the exchange. For example, a one percent change in the network (exchange) impressions induces a 0.09% (0.07%) increase in the exchange (network) conversions in the first phase. This shows that decisions for both channels should be made in tandem after incorporating their interdependencies in the optimization model, and that a failure to do so will result in suboptimal decisions, i.e., possibly underspending on some actions. Note also that the exchange conversions were higher in phases 2 and 3 than those of phase 1, while network conversions showed a decrease in phases 2 and 3 compared to phase 1. While it is important to note that there are significant differences in the number of conversions across phases, we need to take a closer look on the conversion rates and/or profit levels to determine which phase was performed best. We will discuss this in later sections.

⟨⟨Table 3.2 about here⟩⟩

3.4.3.1 Validity of Instruments

The estimation procedures outlined in Section 3.4.2.2 is based on the 3SLS technique. In this section, we confirm the validity of the instrumental variables by providing relevant statistics.

Table 3.3 provides the first-stage regression estimation results. The positive and significant coefficients for $\ln\left(p_{E,t}/\widehat{c_{E,t}}\right)$ and $\ln\left(p_{N,t}/\widehat{c_{N,t}}\right)$ indicate that the number of impressions used in the two channels are positively affected by the respective price-cost ratios. We also observe that both of the dummy variables for Phase 2 and 3 positively (negatively) affect the exchange (network) impressions. We observe this pattern from the 3SLS estimation results provided in Table 3.2, i.e., the impressions and conversions move in the same direction across different phases, which again underscores the importance of looking at the profit level of the campaign (revenue from realized conversions less the cost of impressions). Note also that the adjusted R^2 values of the first stage regressions are high (0.71 and 0.65), which shows that the instrumental variables have significant explanatory power and justifies the relevance conditions for the instrumental variables. Finally, the F-statistics for the joint significance of the first-stage estimations are both over 10, suggesting that the instruments are not weak ([51]).

⟨⟨Table 3.3 about here⟩⟩

3.5 Simulation Analysis

3.5.1 Single Informed Decision Maker

In this section, we explore a single informed decision maker's optimal decisions had he/she recognized and fully exploited the complementarity effects between the two channels in all three phases of the campaign, and compare the results with the status quo. We use the PPA and CPM in each channel to generate decisions, conversions, and profits to compare the performances of each scheme.

We first describe the simulation procedures. Using the impression elasticities for the conversions from Section 3.4.3, we generate the agency's decisions, $x_{E,t}$ and $x_{N,t}$, through Equations (3.3) and (3.4). These decisions represent the agency's decisions had it fully acknowledged complementarities between channels and incorporated them into their optimization problem. Plugging these decisions back into the conversion functions (3.1) and (3.2), we can generate the number of conversions from both channels and calculate the profit levels for the agency, the network, and the overall supply chain. We find that the simulated optimal decisions by the informed decision maker as well as the number of resulting conversions are higher than the actual numbers from the data, and that the agency's profit level in this simulated case is higher than that of the status quo in all three phases. Table 3.4 provides the mean values of the impressions, conversions, and the agency's profit as compared to the status quo. We can see, for example, that the agency's profit in the single informed decision maker scenario for Phase 1 increases by 32% and the operating levels for the agency more than doubles for both the exchange and the network. Acknowledging complementarities between channels clearly boosts the profit levels of the agency and the supply chain. That is, the agency can exploit its informational advantage in better allocating its budget to both channels not only to achieve higher profit but also to increase that of the other players in the supply chain.

⟨⟨Table 3.4 about here⟩⟩

We further explore the agency's decisions in the three phases. Specifically, we investigate how clustered the agency's decisions are by comparing (partial) correlation between the agency's decisions and cosine similarities in the actual and the simulated scenarios.[4] It is well known that complementarities imply that researchers should be able to observe two phenomena: (1) the clustering of business practices, which is manifested through (partial) correlations and (2) the simultaneous presence of the complementary practices affecting business performance more than the sum of individual effects. We verified the latter through IV estimations in the previous section. To better analyze the decision makers' capability of understanding and incorporating the

[4]Cosine similarity measures similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them. Mathematically, given two vectors A and B, the cosine similarity, $\cos\theta$, is represented using the following formula: $\cos\theta = \frac{A \cdot B}{||A||_2 ||B||_2}$.

effect of complementarities into their decision making process, we should also investigate how clustered the business practices are in all three phases. If the decision makers (managers in the agency in our case) acknowledge and exploit complementarities, we would expect that they adopt the practices jointly leading to high correlations and cosine similarity measures. Table 3.5 summarizes partial correlations between the two agency's decisions, exchange and network impressions, in the actual and simulated scenarios, and cosine similarities between the actual decisions and simulated decisions. All correlations include controls for differences in the price cost ratio between the two channels.

⟨⟨Table 3.5 about here⟩⟩

We find that in all three phases, the partial correlations between the exchange and network impressions in the actual scenario are less than those in the simulated single informed decision maker scenario. For example, in Phase 3, the partial correlation in the optimal case is 0.96 whereas it is less than half of that in the actual case (0.40). We also verify that the correlations are higher in the centralized cases (Phases 1 and 3) than that of the decentralized (Phase 2), which leads us to conclude that the centralized decision making structure is better in recognizing and exploiting the effect of complementarities between the exchange and network channels. Lastly, Manager A (0.48) is more capable of embracing the effect of complementarities than Manager B (0.40).

The cosine similarity results also lead us to similar conclusions. That is, the decision vectors in the centralized modes (Phase 1 and 3) are more similar to the respective single informed decision maker's decision vectors than is the decision vector in the decentralized mode (Phase 2). Specifically, the angle between the optimal vector and the actual vector in Phase 1 is 12° while it is 32° in Phase 2. This reconfirms the observation that the centralized mode better exploits the effect of complementarities than the decentralized mode.

3.5.2 Informed Dual Decision Maker Scenario

We now explore a different scenario where there are two informed decision makers, one in charge of the exchange and the other in charge of the network. The decision makers acknowledge the effect of potential spillovers from the other channel but makes locally optimal decisions by maximizing its own channel's profits. Specifically, the exchange manager's expected profit maximization problem is defined as

$$\max_{x_{E,t}} \pi_E = p_{E,t}\, \mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right] - c_{E,t}x_{E,t},$$

while the network manager's problem is defined as

$$\max_{x_{N,t}} \pi_N = p_{N,t}\, \mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right] - c_{N,t}x_{N,t}$$

and their optimal decisions are give as

$$x_{E,t} = \left(A_E \alpha_E \left(\frac{p_{E,t}}{c_{E,t}}\right) x_{N,t}^{\beta_E} e^{\frac{(\sigma_E)^2}{2}}\right)^{\frac{1}{(1-\alpha_E)}} \tag{3.11}$$

$$x_{N,t} = \left(A_N \alpha_N \left(\frac{p_{N,t}}{c_{N,t}}\right) x_{E,t}^{\beta_N} e^{\frac{(\sigma_N)^2}{2}}\right)^{\frac{1}{(1-\alpha_N)}} \tag{3.12}$$

The simulation procedure is similar to that in Section 3.5.1. We first generate the agency's decisions as given in Equations (3.11) and (3.12). We plug these decisions on impressions to the conversion functions (3.1) and (3.2) to generate the conversions and calculate the profit levels of for the agency, the network, and the supply chain. Table 3.4 reports the mean values of the impressions, conversions, and the agency's profit as compared to the status quo. In Phase 1, for example, the dual informed decision maker's scenario can achieve 24% higher agency's profit than what the original data gives us. It is still 8% less than the single informed decision maker scenario. This pattern persists in all other phases, which suggests that the large chunk of increase in profit comes from the decision maker's recognition of complementarities between channels. A similar conclusion can be made by observing the partial correlation values between the exchange and network impressions and cosine similarity values between the actual decision vectors and the simulated ones. In Phase 3, the correlation between the exchange and network impressions in the dual informed decision maker scenario is 0.93, which is considerably higher than that of the status quo (0.40) but is slightly less than that of the single informed decision maker scenario (0.96).

3.5.3 Interrupted Time Series Analysis Framework and Results

When considering the impact of an intervention or policy change with no control group as in our case, an interrupted time series analysis (ITS) can be used given that there are multiple observations on an outcome variable of interest in the pre- and post-intervention periods. The ITS approach offers a quasi-experimental research design with a high degree of internal validity (see

[13] and [17]). As statistical analyses used for ITS must account for autocorrelation in the data, we use autoregressive integrated moving-average (ARIMA) based dynamic regression models. The standard ITS dynamic regression model with ARIMA(p, d, q) assumes the following form:

$$y_t = \phi_0 + \beta_1 T + \beta_2 x_t + \beta_3 T x_t + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q} + \epsilon_t \tag{3.13}$$

where T is the time elapsed, x_t is a dummy variable indicating pre-intervention (0) or post-intervention (1), and y_t is the outcome variable at time t, differenced d times. Notice the predictors on the right hand side include both p lagged values of y_t and q such ϵ_t. The coefficient β_1 represents the change in the outcome variable with respect to a unit increase in time (which shows the underlying pre-intervention trend), β_2 indicates the level change after the intervention (compared to the counterfactual), and β_3 is interpreted as the slope change after the intervention.

The key outcome variables that we are focused on are the following: (1) the angle between the status quo and simulated decision vectors, (2) the proportion of the total status quo impressions to the total simulated impressions, and finally (3) the proportion of status quo profit to the simulated profit. The ITS analysis procedure is described below.

Step 1: Exclude outliers that are more than 1.5 interquantile ranges (IQRs) below the first quantile or above the third quantile from the outcome variables.

Step 2: Fit an ARIMA-based dynamic regression model on the data from Step 1 and use the Kalman filter ([26]) to replace the excluded values.[5]

Step 3: Re-fit an ARIMA-based dynamic regression model on the data from Step 2.

3.5.3.1 Angle between Status Quo and Simulated Decision Vectors

The dynamic regression model for the angles between the status quo and single informed decision maker's decision vectors is an ARIMA(1,0,1) model:

$$\begin{aligned} y_t = \underbrace{31.0803}_{1.4038} - \underbrace{22.8396}_{1.9693} * Phase\ 1 - \underbrace{20.7880}_{1.9388} * Phase\ 3 + \underbrace{0.7987}_{0.0847} y_{t-1} \\ - \underbrace{0.4449}_{0.1202} \epsilon_{t-1} + \epsilon_t \end{aligned}$$

Standard error for each coefficient is reported underneath the underbrace. We excluded the time variable and the interaction term between the time and dummy variables because they were not significant. Moreover, the above model led to better fit in terms of AICc values. We can observe that the angle between the status quo and simulated decision vectors are significantly lower in Phases 1 and 3 than those in Phase 2 by 22.84 and 20.79 respectively, which implies that the agency tends to get the closer-to-optimal mix of impressions when there is a centralized decision maker. We also conduct

[5] We can also use Kalman smoother. Dealing with missing values using the Kalman filter is interpreted as extrapolation of the series while using the Kalman smoother is interpreted as interpolation of the observed series.

the Ljung-Box test for autocorrelation which yields a p-value of 0.91 so we cannot reject the null hypothesis that the model does not exhibit lack of fit. By design, the (single-group) ITS analysis has no comparable control group. Hence, the pre-intervention line or trend can be projected into the treatment period and serve as the counterfactual. Figure 3.1 visualizes the estimation results and counterfactuals. The solid red lines on the graph represent the predicted values in each phase based on the dynamic regression model, while the dashed red lines represent the counterfactuals. For example, the dashed red line in Phase 2 can be interpreted as the angle between the status quo and simulated decision vectors if Manager A from Phase 1 continued to manage both channels through Phase 2. The grey and black dashed lines represent the 80% and 95% prediction intervals, respectively. We can observe that the counterfactuals lead to significantly different results than the predicted values in each phase (the solid red lines lie outside the prediction intervals), i.e., there are significant level changes across different phases.

⟨⟨Figure 3.1 about here⟩⟩

Similarly, we can compare the status quo decision vectors to those from the dual informed decision maker scenario. The ITS analysis yields the following ARIMA(1,0,1) model:

$$\begin{aligned} y_t = \underbrace{35.0904}_{1.6526} - \underbrace{29.7370}_{2.2604} * Phase\ 1 - \underbrace{25.7599}_{2.3343} * Phase\ 3 + \underbrace{0.7595}_{0.0838} y_{t-1} \\ - \underbrace{0.3078}_{0.1188} \epsilon_{t-1} + \epsilon_t \end{aligned}$$

Figure 3.3 illustrates the results. Qualitatively, this result is similar to that reported in the single informed manager case.

⟨⟨Figure 3.3 about here⟩⟩

3.5.3.2 Proportion of Status Quo Impressions to Simulated Impressions

Next we analyze the total spend on the impressions, i.e., we study how the total advertising efforts change over time and across phases as compared to the simulated scenarios. The ITS analysis on the proportion of the status quo impressions to the single informed manager's impressions yields the following AR(1) dynamic regressions model:

$$y_t = \underbrace{0.2942}_{0.0311} + \underbrace{0.2028}_{0.0449} * Phase\ 1 + \underbrace{0.1234}_{0.0442} * Phase\ 3 + \underbrace{0.3610}_{0.0697} y_{t-1} + \epsilon_t$$

The Ljung-Box test for autocorrelation yields a p-value of 0.17. Notice that the dependent variables are significantly higher in Phases 1 and 3 than those in Phase 2 by 0.20 and 0.12 respectively, which suggests the total spend in advertising in the decentralized case is considerably lower than that in the centralized scenario. As in Section 3.5.3.1, we project the pre-intervention line into the treatment period to study the counterfactuals (see Figure 3.5). It is important to note that the predicted lines in each phase (the solid red lines) are inside the counterfactual prediction intervals (the black and grey dashed lines). This suggests that if Manager A continued to managed the campaign alone through Phase 2, he/she would not exert significantly different advertising effort than the status quo. Similarly, if the two managers in Phase 2 continued to manage the campaign through Phase 3, their total spend in impressions would not be significantly different from the status quo. This suggests that the total spend in impressions is not impacted by the organizational design or decision making structure of the agency in a counterfactual sense. The analysis on the dual decision maker scenario leads to similar results (see Figure 3.7).

⟨⟨Figures 3.5 and 3.7 about here⟩⟩

3.5.3.3 Proportion of Status Quo Profit to Simulated Profit

In this section, we analyze the proportion of status quo profits to simulated profits. The ITS analysis gives us the following regression model:

$$y_t = \underbrace{0.1588}_{0.1022} + \underbrace{0.0029}_{0.0011} T + \underbrace{0.5220}_{0.0803} * Phase\ 1 - \underbrace{0.0045}_{0.0813} * Phase\ 3 + \epsilon_t$$

Notice in this model we have a positive trend line as well, which suggests that as time progresses the agency becomes more cost-effective. It is also important to note that the coefficient for *Phase* 3 is not significant, i.e., there is no significance level change from Phase 2 to Phase 3. A counterfactual analysis yields Figure 3.9. There is a key difference between this counterfactual result with that in Section 3.5.3.2. Recall that the total spend in impressions was not impacted by the organizational structure in a counterfactual sense. In other words, when we projected the pre-intervention line into the treatment period, the counterfactual prediction intervals included the predicted lines in both Phases 2 and 3. However, notice from Figure 3.9 that it is not the case here in Phase 2. The solid red line lies outside the counterfactual prediction interval. This implies that the there clearly is a level change from Phase 1 to Phase 2 in a counterfactual sense. Notice also that this does not hold in Phase 3, i.e., there is no significant level change from Phase 2 to Phase 3. Analyzing the dual informed managers' scenario leads to qualitatively similar results (see Figure 3.11). From anecdotal evidence and historical data, we learned that Manager A has much more experience in carrying out ad campaigns in terms of both how long he/she has been working and how many ad campaigns he/she has dealt with. Thus our results imply that while the organizational design has an impact on the profit levels as compared to those from the simulated scenarios, experience of a manager also plays a big role in either strengthening or weakening the magnitude of the impact.

⟨⟨Figures 3.9 and 3.11 about here⟩⟩

3.5.3.4 Robustness Checks

To improve the robustness of our analysis, we address a couple of distinctive issues with the time series data. First, we check if there is any seasonal pattern in the dependent variables discussed in Sections 3.5.3.1 through 3.5.3.3. We used a range of functions such as Fourier terms (pairs of sine and cosine functions) or splines, but we found that the association between the intervention and the outcome variables are largely unaffected. Second, we test whether our empirical model captures a causal effect in periods where there is no intervention. In other words, we test whether or not we are able to observe a continuous trend or line in the absence of an intervention. We look to the regression-discontinuity literature (see [28]) for this analysis. In implementing this robustness test, we test for interruptions by replacing the true intervention with other pseudo-interventions in the pre-intervention time period. We adopt a simple iterative process of testing each pre-intervention time periods as the pseudo-intervention. We did not find any statistically significant estimates from these robustness tests with pseudo-interventions, which provides strong support that our empirical framework presents causal impact of the change in organizational designs. Lastly, we add more relevant control variates to our empirial specification to check whether the prediction intervals in, for example, Figures 3.5 and 3.9 are robust to different specifications. Specifically, we add the exchange and network CPMs, $c_{E,t}$ and $c_{N,t}$, which are important factors driving the managers' decisions to the general ARIMA model in Equation (3.13) and rerun the ITS analysis (see Figures 3.2, 3.4, 3.6, 3.8, 3.10, and 3.12). We find that our results from Sections 3.5.3.1 through 3.5.3.3 are qualitatively similar to this specification.

3.6 Conclusion

In this paper, we investigated how the organizational design of a firm can affect its ability to recognize and exploit the cross-channel complementarities in the context of online advertising. To the best of our knowledge, our study takes the first step of empirically verifying whether a firm's decision making structure makes a difference in the performance effects of complementarities. Through our unique dataset which provided a natural experimental setting with three different phases of organizational design, we initially focused on estimating the spillover effects across two different channels using a iterative 3SLS approach. After establishing that complementarities are present, we then explore the effect of centralization versus decentralization on recognizing and exploiting the synergy effects by using simulation and counterfactual analysis. We find that the centralized mode, with a single manager for both channels, better recognizes and incorporates complementarities between the two channels into its decision making, while the decentralized mode with two managers, one in charge of each channel, systematically fails in allocating the right amount of impressions in each channel, leading to less profit for the agency. Our analysis demonstrates that the decision structure can make a difference in the performance due to more effective exploitation of cross-channel synergies in an online advertising campaign.

The key implication of our results involves the need to train decentralized decision makers to recognize and act upon complementarity (when present), and to even consider reorganizing the decision making structure to exploit such synergies. More specifically, our results underscore the need to redesign the agency's current organizational structure from specializing in media buying (the status quo) to campaign based management, where one (or more) decision maker(s) in charge of a campaign is (are) made aware of horizontal synergies as well as informational aspects involved in the supply chain. The status quo has separate and specialized groups of managers for the exchange and the network. The current organizational structure reflects a rationalization of the role of the agency in the supply chain based on economies of scale in media buying and lowering search and transaction costs for the advertisers; it is well-suited for exploring new publishers and brokers in the two channels, and having access to their online content. However, our results suggest a broader rationalization of the economic role of the agency, one which provides a significant increase in profit potential by considering cross-channel interactions when making decisions for each channel. Hence, we propose that the agency group together managers in charge of similar campaigns, and either integrate the network and the exchange management teams or ensure through interventions that they understand the impact of synergies on their decision making and agency profits. Such structural changes will involve additional costs of cross-training managers as well as business process changes to move from a silo mode of operation to a more integrated focus on the overall supply chain; nevertheless, our numerical results also show that there is a significant upside potential from such an initiative.

Our findings also show that setting the right incentive system is also key to improved performance. The current incentive system in the agency is based on the number of realized conversions each manager extracts from his/her campaign. Naturally, the channel managers tend to put less emphasis on the cost of impressions and even less on the potential interactions across channels. A gainsharing system for each campaign may motivate the managers to better work as a team and improve their campaign performances.

3.7 Tables and Figures

	Total	Exchange Total	Network Total	Exchange Daily Average	Network Daily Average
Phase 1					
Mille Impressions	279,682	235,268	44,414	4,127	779
Conversions	14,355	8,062	6,293	141	110
Profit	30,651	13,804	16,847	242	296
Pay Per Action				5.00	5.01
Cost Per Mille				0.16	0.50
Conversion Rate (%)	0.0051	0.0034	0.0142		
Phase 2					
Mille Impressions	229,843	213,274	16,569	3,385	263
Conversions	12,567	9,404	3,163	149	50
Profit	25,095	13,837	11,258	220	179
Pay Per Action				5.20	5.25
Cost Per Mille				0.17	0.37
Conversion Rate (%)	0.0055	0.0044	0.0191		
Phase 3					
Mille Impressions	947,212	860,010	87,202	14,333	1,453
Conversions	31,490	21,361	10,129	356	169
Profit	61,176	34,177	26,999	570	450
Pay Per Action				5.25	5.26
Cost Per Mille				0.09	0.34
Conversion Rate (%)	0.0033	0.0025	0.0116		

Table 3.1: Descriptive Statistics

Variables	$\log y_{E,t}$	$\log y_{N,t}$
Intercept	-0.8307***	-0.5080**
$\log x_{E,t}$	0.6413***	0.0703*
$\log x_{N,t}$	0.0954***	0.6972***
Phase 2	0.4396***	-0.1586˙
Phase 3	0.4591**	-0.3183*
Trend	-0.0037***	0.0015
Adjusted R^2	0.9156	0.9366
McElroy R^2	0.9340	

$^{\cdot}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 3.2: IV Estimation Results

Variables	$\ln x_{E,t}$	$\ln x_{N,t}$
Intercept	-0.9792	-2.5089^{***}
$\ln p_{E,t}/\widehat{c_{E,t}}$	2.2651***	-
$\ln p_{N,t}/\widehat{c_{N,t}}$	-	2.2206***
Phase 2	0.8843***	-1.5537^{***}
Phase 3	1.4677***	-0.8699^{*}
Adjusted R^2	0.7121	0.6505

$^{\cdot}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 3.3: First-Stage Regressions of Impression Variables

	Exchange Impressions (%)	Network Impressions (%)	Exchange Conversions (%)	Network Conversions (%)	Agency's Profit (%)
Phase 1					
Actual	100	100	100	100	100
Informed Dual	143	244	146	171	124
Informed Single	220	425	202	263	132
Phase 2					
Actual	100	100	100	100	100
Informed Dual	110	578	118	259	167
Informed Single	192	617	174	313	177
Phase 3					
Actual	100	100	100	100	100
Informed Dual	147	370	150	264	159
Informed Single	220	633	203	396	170

Table 3.4: Actual versus Other Scenarios

	Partial Correlations	Average Cosine Similarity
Phase 1		
Actual	0.48^{***}	0.9780 (12.03)
Informed Dual	0.92^{***}	0.9746 (12.94)
Informed Single	0.95^{***}	-
Phase 2		
Actual	0.12	0.8491 (31.88)
Informed Dual	0.84^{***}	0.8053 (36.36)
Informed Single	0.91^{***}	-
Phase 3		
Actual	0.40^{**}	0.9813 (11.09)
Informed Dual	0.93^{***}	0.9846 (10.08)
Informed Single	0.96^{***}	-

Note. Partial correlations controlling for differences in price cost ratio between the Exchange and Network. Test is against the null hypothesis that the correlation is zero. The numbers in parentheses in the cosine similarity column represent degree of angles.

$^{\cdot}p < 0.1$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 3.5: Partial Correlations and Cosine Similarities

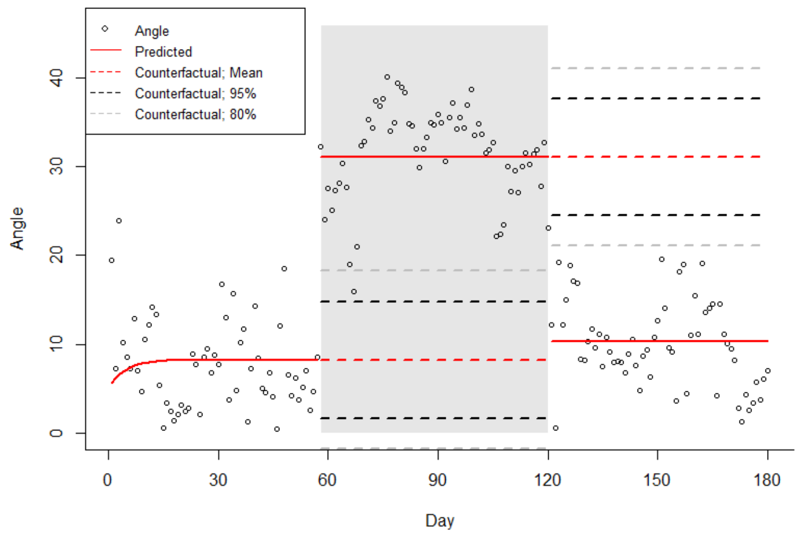

Figure 3.1: Angle Between Actual and Single Informed Manager's Impressions

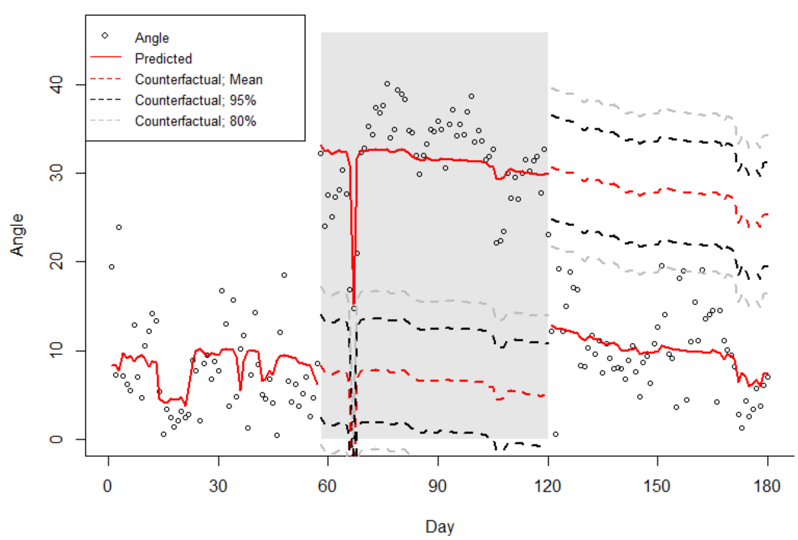

Figure 3.2: Angle Between Actual and Single Informed Manager's Impressions

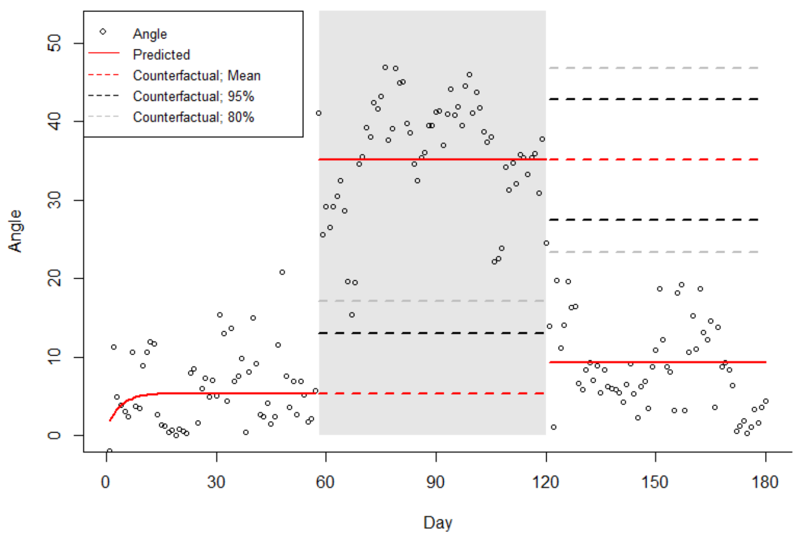

Figure 3.3: Angle Between Actual and Dual Informed Managers' Impressions

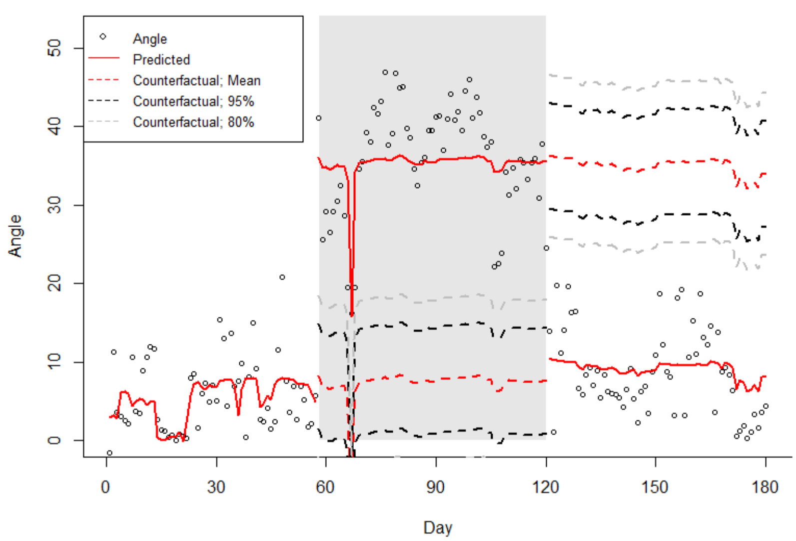

Figure 3.4: Angle Between Actual and Dual Informed Managers' Impressions

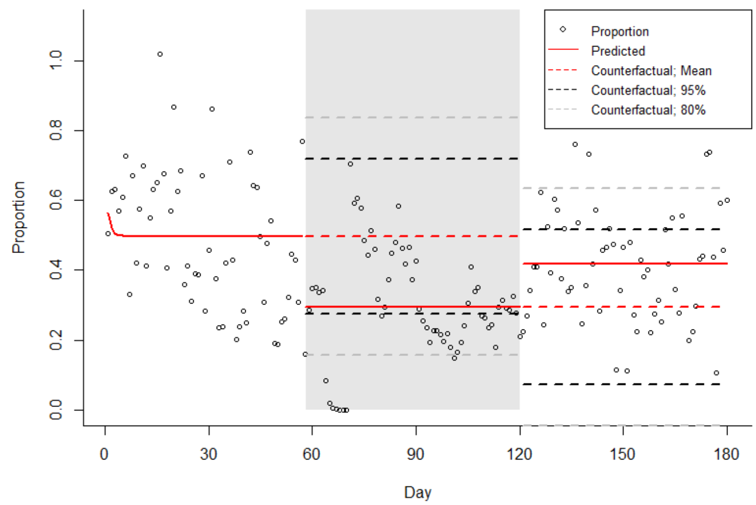

Figure 3.5: Proportion of Actual Impressions to Single Informed Manager's Impressions

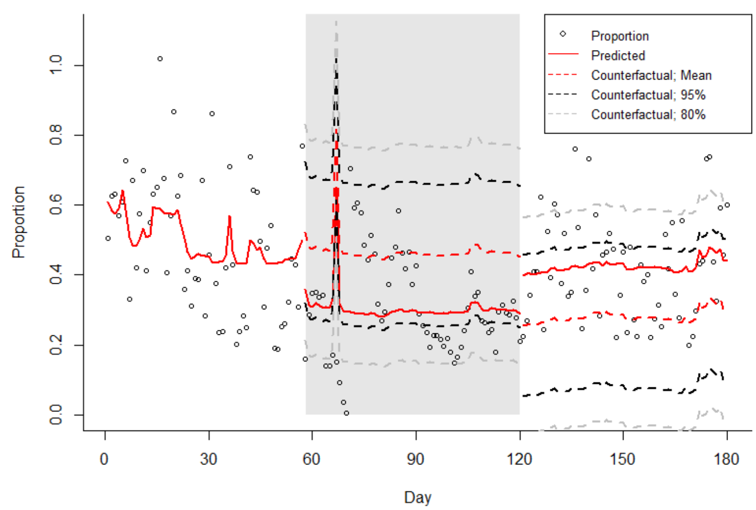

Figure 3.6: Proportion of Actual Impressions to Single Informed Manager's Impressions

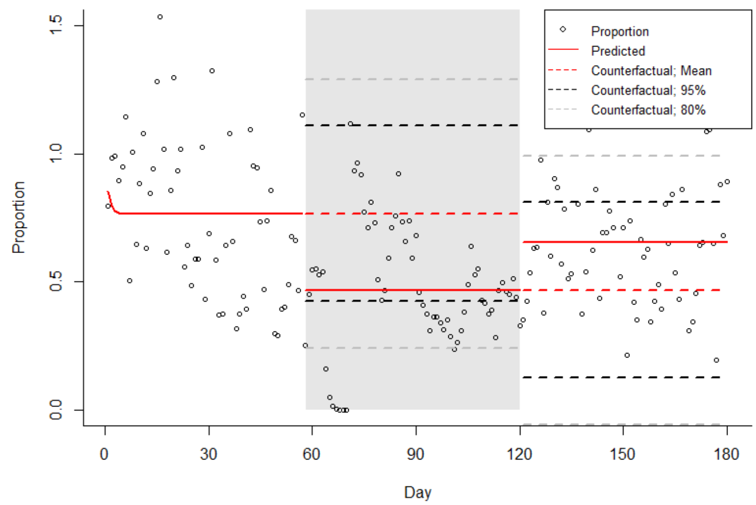

Figure 3.7: Proportion of Actual Impressions to Dual Informed Managers' Impressions

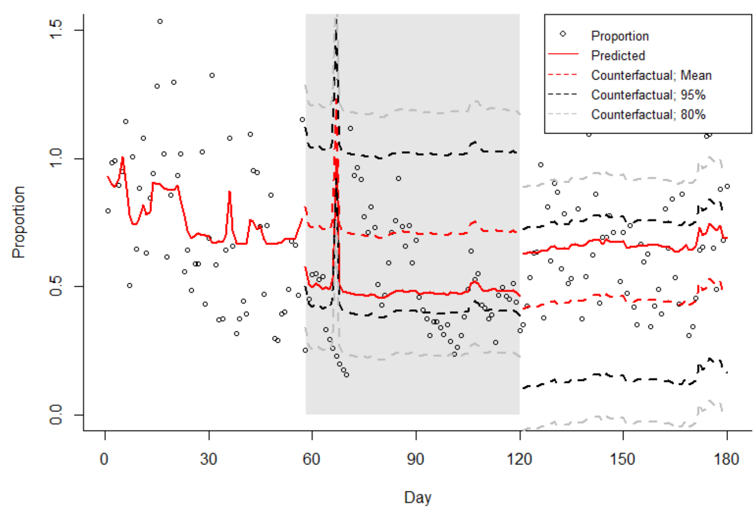

Figure 3.8: Proportion of Actual Impressions to Dual Informed Managers' Impressions

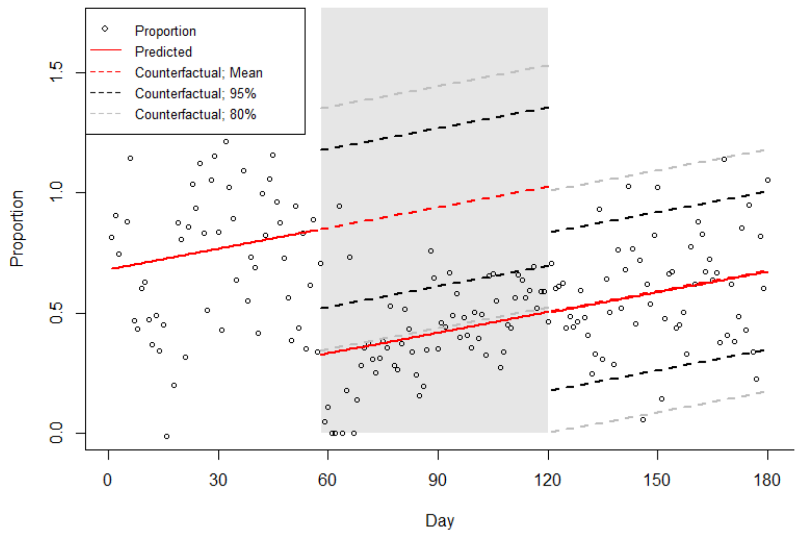

Figure 3.9: Proportion of Actual Profit to Single Informed Manager's Profit

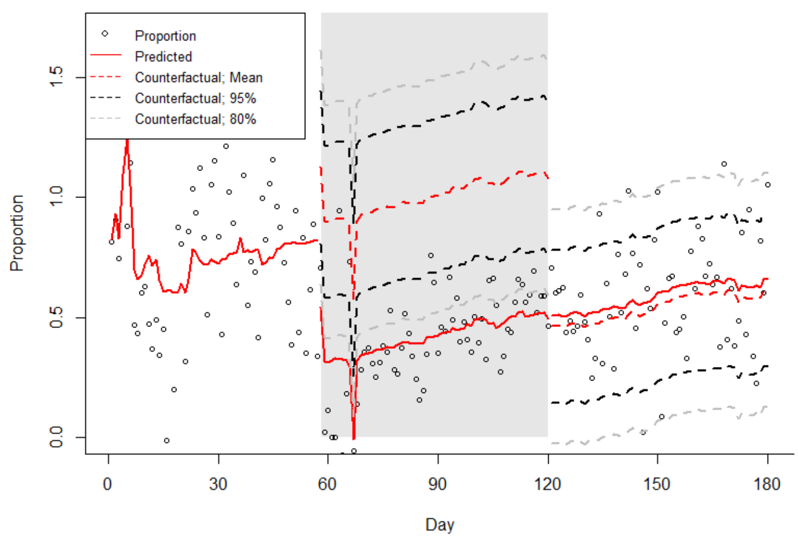

Figure 3.10: Proportion of Actual Profit to Single Informed Manager's Profit

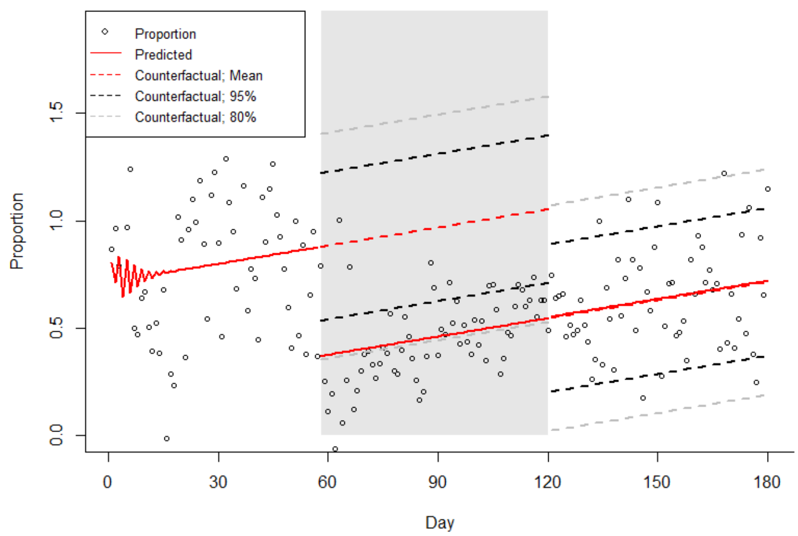

Figure 3.11: Proportion of Actual Profit to Dual Informed Managers' Profit

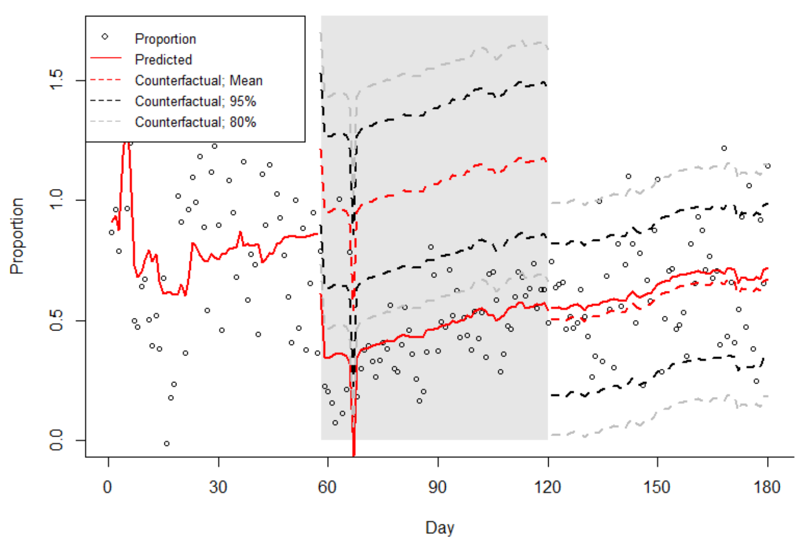

Figure 3.12: Proportion of Actual Profit to Dual Informed Managers' Profit

Chapter 4

On the Challenges of Detecting Complementarities in a Structural Model

4.1 Introduction

We analyze complementarities in a dual-channel online advertising supply chain consisting of an advertiser, agency, publishers and brokers, and estimate their magnitudes using two distinct types of models: (1) one based on instrumental variables (IV)[1], and (2) an explicit *structural model*[2] that accounts for a decision maker's optimizing behavior. The central question in our research is how operating decisions accounting for (or ignoring) complementarities can affect the empirical evidence and the ability of structural models to correctly measure such effects. Specifically, we demonstrate that if a decision maker (in our specific setting, a manager in the agency, who decides on the number of impressions to use in both channels) ignores complementarities (possibly due to a variety of reasons, including lack of cross-channel training, decentralized structure, etc.), an explicit structural model systematically

[1]As emphasized in [4], an IV framework enables us to disentangle complementarity from clustered organizational practices.

[2]This model has been referred to as a more structural model by, for example, Brynjolfsson and Milgrom (2012). This involves simultaneous estimation of a system of equations with potential cross-equation restrictions.

underestimates the complementarity effect, and can make true *complements* appear to be *substitutes*, even when the decision maker's behavior is correctly reflected in the structural model. The IV approach, however, will correctly identify the effects regardless of the decision maker's ability to recognize and act on complementarities. This result is driven by the fact that in an explicit structural model, suboptimal decisions force the researcher to make one of two bad choices: either (1) make unwarranted assumption of the decision maker acting on complementarity (when in reality the decisions are being made naively) or (2) reflect the decision making deficiencies in the model, leading in both cases to misspecification biases. By contrast, the IV model is agnostic with respect to managerial behavior, and thus produces more robust estimates under these circumstances. The implications of our study are that the researcher needs to have an in-depth understanding of the decision process itself before developing a structural model. If s/he is unsure about whether the decision maker recognizes complementarities, or is confident that such interactions are being ignored, the IV approach will provide more reliable estimates of any potential complementarities than the structural model, which may cause a significant bias in the estimates. An important caveat, however, is that the IV approach can sometimes be highly inefficient relative to the explicit structural model. Hence the researcher needs to introduce sufficient relevant covariates and instrumental variables such that the variables adequately explain the variations in the dependent variables.

In the extant empirical literature on complementarity in economics (e.g., [3], [4], [39]) and information systems (e.g., [8], [62]), no distinction has been made between (1) the existence of complementarities in organizational practices and (2) whether decision makers recognize and act upon such complementarities. To the best of our knowledge, our study is the first attempt to distinguish between the measurement of complementarities when decision makers act or fail to act on them.

We study an online advertising supply chain, which involves an online advertising agency who attempts to sell a product using two different channels; a network of large publishers, and an exchange with real-time bidding for impressions on publishers' websites. The advertising network is a closed group of agencies and publishers akin to a privately traded market, where prices are determined through individual negotiations, while the advertising exchange is a technology-driven platform that facilitates buying and selling of impressions inventory with prices determined by a real-time-bidding (RTB) algorithm. The agency buys impressions from the two channels, and a conversion is realized when a customer carries out a measurable action, e.g., a customer makes a purchase, provides information, or downloads an app, in response to an impression. Then the agency is paid a price for each conversion from an advertiser, and the agency's profit margin is the difference between the revenue they receive from the advertiser and the cost they pay for impressions spent on both channels. Conversions in the two channels may not be independent as the impressions are targeted at the same set of customers. However, it is not clear if there are synergies present or if there is a substitution effect. That is, we do not know if increasing impressions in one channel will improve or hinder the effectiveness of the impressions in the other channel.

4.2 The Online Advertising Supply Chain

In the online advertising campaign we study, the supply chain consists of an ad agency, a network in which the agency runs ads on a single publisher, and an exchange in which the agency buys impressions from a single broker through real time bidding. At the beginning of a campaign, an advertiser and the agency negotiate the pay per action (PPA)[3] for the exchange and the network, $p_{E,t}$ and $p_{N,t}$, the prices paid by the advertiser to the agency per each conversion. These prices can vary over the course of a campaign through renegotiation between the advertiser and the agency, and can also be different for conversions from the network and the exchange, though such variation appears to be minimal in our data. Once the PPAs are set, the agency decides the number of impressions, $x_{E,t}$ and $x_{N,t}$, it wishes to spend in the exchange and the network, respectively, given the cost per mille (CPM)[4] in each channel. Then, conversions, $y_{E,t}$ and $y_{N,t}$, are realized in both channels, and the agency is paid by the advertiser for all *realized* conversions in the network and the exchange.

[3]PPA is the amount of money that the agency receives from an advertiser every time an advertisement leads to a specified action (e.g., a sale, click, download, or subscription).

[4]The underlying assumption for the exchange is that every participant in the exchange is a price taker. Moreover, the agency has a fair assessment of the CPM through extensive learning before the launching of a campaign.

Multiple studies (e.g., [1], [31], [57], etc.) provide theoretical and empirical support for interactions across channels. If the agency considers the inter-channel interactions in its decisions (we will refer to this scenario as the *informed agency*), its view of the conversion process in a channel in its simplest form will include impressions from the other channel as

$$y_{E,t} = A_E x_{E,t}^{\alpha_E} x_{N,t}^{\beta_E} e^{u_E} \tag{4.1}$$

$$y_{N,t} = A_N x_{N,t}^{\alpha_N} x_{E,t}^{\beta_N} e^{u_N} \tag{4.2}$$

where u_E and u_N are normally distributed with zero means, variances σ_N^2 and σ_E^2, and possibly nonzero covariances. The agency's expected profit maximization problem can be written as

$$\max_{x_{E,t}, x_{N,t}} \pi_A = p_{E,t}\,\mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right] - c_{E,t}x_{E,t} + p_{N,t}\,\mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right] - c_{N,t}x_{N,t}. \tag{PM}$$

where $\mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right] = A_E x_{E,t}^{\alpha_E} x_{N,t}^{\beta_E} e^{\frac{\sigma_E^2}{2}}$, $\mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right] = A_N x_{N,t}^{\alpha_N} x_{E,t}^{\beta_N} e^{\frac{\sigma_N^2}{2}}$. Maximizing the expected profit with these conversion processes involves finding solutions to a nonlinear system of equations given below:

$$\frac{\partial \pi_A}{\partial \pi_E} = p_{E,t}\frac{\partial\,\mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right]}{\partial x_{E,t}} - c_{E,t} + p_{N,t}\frac{\partial\,\mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right]}{\partial x_{E,t}} = 0 \tag{4.3}$$

$$\frac{\partial \pi_A}{\partial \pi_N} = p_{E,t}\frac{\partial\,\mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right]}{\partial x_{N,t}} + p_{N,t}\frac{\partial\,\mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right]}{\partial x_{N,t}} - c_{N,t} = 0 \tag{4.4}$$

We can estimate Equations (4.1) through (4.4) when the researcher believes that the agency is informed.

On the other hand, although managers in the agency technically have full visibility of decisions made in both the network and the exchange through the data management platform, we observed a decentralized structure of decision making and the complete lack of cross-channel communication and coordination within the agency, as a result of which the managers did not consider the interplay between the two channels in their decisions. Thus, while ideally the agency's perspective of the conversion processes in each channel should be functions of impressions in both the network and the exchange (as shown in Equations (4.1) and (4.2)), to model the decision process we actually witnessed, we can define the agency's *naive* view of the conversion processes in each channel in their simplest forms as

$$y_{E,t} = A_E^M x_{E,t}^{\alpha_E^M} e^{u_E^M} \tag{4.5}$$

$$y_{N,t} = A_N^M x_{N,t}^{\alpha_N^M} e^{u_N^M} \tag{4.6}$$

where u_N^M and u_E^M are normally distributed with zero means and variances $\left(\sigma_N^M\right)^2$ and $\left(\sigma_E^M\right)^2$ (and possibly nonzero covariances), and where the superscript M represents the naive agency. The maximizing condition then yields the optimal numbers of impressions, $x_{E,t}^M$ and $x_{N,t}^M$, which are functions of the price-cost ratio in their respective channels:

$$x_{E,t}^M = \left(A_E^M \alpha_E^M \left(\frac{p_{E,t}}{c_{E,t}} \right) e^{\frac{\left(\sigma_E^M\right)^2}{2}} \right)^{\frac{1}{\left(1-\alpha_E^M\right)}} \tag{4.7}$$

$$x_{N,t}^M = \left(A_N^M \alpha_N^M \left(\frac{p_{N,t}}{c_{N,t}} \right) e^{\frac{\left(\sigma_N^M\right)^2}{2}} \right)^{\frac{1}{\left(1-\alpha_N^M\right)}} \tag{4.8}$$

We will use these insights from the naive agency's problem in choosing our instruments in our IV analyses in Section 4.3. Moreover, in later sections we estimate Equations (5) through (8) when the researcher believes that the agency is naive.

4.3 Data Description and Instrumental Variables Analyses

We use data from an online advertising agency to test spillovers between the exchange and network channels. We study an advertiser who ran online campaigns on both channels through the ad agency for 272 days. The advertiser provides an online service, and a conversion for the advertiser's campaign is a subscription to its website. Table 4.1 provides descriptive statistics for the campaign we study.

⟨⟨Table 4.1 about here⟩⟩

Each observation in our dataset consists of a day of operations. In each row, we have the number of impressions, the number of conversions, PPA values, revenue, cost, and profit levels corresponding to the publisher in the network and the broker in the exchange. We then use this information to further construct additional variables such as the CPM values and conversion rates (CR) for each channel. The typical industry definition of CR is the number of conversions from an ad divided by the number of clicks on the ad, i.e., CR = conversions/clicks. In this paper, however, we will define CR to be the number of conversions divided by the number of impressions. Also, since the actual CPM bids are not available in our analysis, we instead use the average daily

CPM as a proxy for the actual bids.

The estimation process is carried out in two main steps. First, we estimate dynamic regression models (AR(1)) for the CPM values in each channel with respective lagged conversion rate as a regressor and use the fitted values from the models to use as bases for the managers' beliefs about the CPMs in making decisions for how much impressions to spend on any given day. Second, we estimate the conversion functions using a two-stage least squares (2SLS, [2]) approach with instrumental variables constructed from insights we gained from the previous section. We describe each step in more detail below.

4.3.1 Dynamic Regression Model for CPM

When making decisions on how many impressions to spend in each channel, the managers in the agency form a belief about the daily average CPM on a given day based on past history. To mimic this belief forming process, we estimate dynamic regression models for the CPM values and use the fitted values from the models as the agency managers' beliefs. More specifically, we observe that the CPM values are generally sticky, i.e., they largely depend on previous period's CPM, $c_{E,t-1}$ and $c_{N,t-1}$. They also depend on previous period's conversion rates, $CR_{E,t-1}$ and $CR_{N,t-1}$, respectively. The higher the conversion rate in the previous period, the larger the CPM bid by the agency so as to secure advertising slots.

We thus estimate a first-order autoregressive (AR(1)) dynamic regression model for each channel with previous period's conversion rate as a regressor. For the exchange we have

$$c_{E,t} = \theta_0 + \theta_1 CR_{E,t-1} + \eta_t,$$
$$\eta_t = \phi\eta_{t-1} + \epsilon_t$$

where ϕ is the first-order autocorrelation coefficient and ϵ_t is white noise which represents the part of CPM that cannot be explained by this model. We can write a similar equation for the network CPM. The fitted values from these models, $\widehat{c_{E,t}}$ and $\widehat{c_{N,t}}$, represent the agency's belief about the CPMs which are used to make decisions on the impressions.

4.3.2 Estimation of Conversion Functions

We now estimate the conversions from each channel on day t, $y_{E,t}$ and $y_{N,t}$ via 2SLS. After log transformation the conversion functions can be written as follows:

$$\ln y_{E,t} = \ln A_E + \alpha_E \ln x_{E,t} + \beta_E \ln x_{N,t} + \kappa_E Trend + u_{E,t} \tag{4.9}$$

$$\ln y_{N,t} = \ln A_N + \alpha_N \ln x_{N,t} + \beta_N \ln x_{E,t} + \kappa_N Trend + u_{N,t} \tag{4.10}$$

If the sum of the elasticities, α and β, in each equation is less than one, as our empirical analysis will later show, the conversion process exhibits decreasing returns to scale; in this case each incremental amount of online advertising causes a lesser increase in conversions, which may be explained as a result of advertising saturation. The coefficients, β_E and β_N, reflect the inter-channel or horizontal spillover effects, so if the coefficients are positive, then we can infer that there are positive spillover effects from one channel to the other. Finally we include the trend variable to control for trend in our regressions.

Note that in each channel's conversion equation its own impression variable is potentially endogenous. That is, exchange (network) impressions, $x_{E,t}$ ($x_{N,t}$), are correlated with the disturbance, $u_{E,t}$ ($u_{N,t}$), because it is very likely that there are variables that affect both the conversions and impressions in the same channel simultaneously. For example, increase in total traffic on a website may drive both the impressions and conversions to increase. Impression variable from the other channel has less of this concern because it is quite unlikely that there exists a common variable that simultaneously affects, for example, both the exchange side conversions and the network publisher's impressions. Due to the above-mentioned endogeneity issue, the OLS estimators are inconsistent. As emphasized in [3] and [4], we can use an instrumental variables (IV) framework to disentangle complementarity from organizational practices. This approach can provide a consistent estimate of the synergy coefficients if one can find instrumental variables that satisfy relevance and exclusion restriction assumptions ([24]), i.e., they should be correlated with the corresponding endogenous regressor (relevance) and uncorrelated with the error term (exclusion restriction).

Our analyses in Section 4.2 yield a natural set of instrumental variables. In particular, the number of impressions in each channel is a function of the respective price-cost ratios in each channel (see Equations (4.7) and (4.8)). Therefore we use these ratios, $p_{E,t}/\widehat{c_{E,t}}$ and $p_{N,t}/\widehat{c_{N,t}}$, as instruments for the impression variables, $x_{E,t}$ and $x_{N,t}$, respectively. Both of these ratios naturally satisfy the relevance condition.[5] In addition, we would expect that these ratios affect the conversions only through the impressions because the general audience of the ads have no information about these ratios when they are making their decisions to *convert*. With these instrumental variables, we carry out the following 2SLS estimation procedure:

Stage 1: Estimate via ordinary least squares (OLS) the endogenous independent variables, $x_{E,t}$ in Equation (4.9) and $x_{N,t}$ in Equation (4.10), using instrumental variables, i.e., $p_{E,t}/\widehat{c_{E,t}}$ for $x_{E,t}$ and $p_{N,t}/\widehat{c_{N,t}}$ for $x_{N,t}$, as well as the trend variable in Equations (4.9) and (4.10). We then calculate the predicted endogenous independent variables, $\widehat{\ln x_{E,t}}$ and $\widehat{\ln x_{N,t}}$.

$$\ln x_{E,t} = \rho_E + \lambda_E \ln\left(p_{E,t}/\widehat{c_{E,t}}\right) + \mu_E Trend + \nu_{E,t}$$

$$\ln x_{N,t} = \rho_N + \lambda_N \ln\left(p_{N,t}/\widehat{c_{N,t}}\right) + \mu_N Trend + \nu_{N,t}$$

Stage 2: Using the predicted endogenous variables from Stage 1, we estimate the coefficients in Equations (4.9) and (4.10) via OLS.

$$\ln y_{E,t} = \ln A_E + \alpha_E \widehat{\ln x_{E,t}} + \beta_E \ln x_{N,t} + \kappa_E Trend + u_{E,t}$$

$$\ln y_{N,t} = \ln A_N + \alpha_N \widehat{\ln x_{N,t}} + \beta_N \ln x_{E,t} + \kappa_N Trend + u_{N,t}$$

Table 4.2 summarizes the key results from the IV estimation, which indicate that there are significant synergy effects between the network and the exchange. A one percent increase in the network (exchange) impressions induces

[5] There of course exists the assumption that the agency is a profit maximizer.

a 0.05% (0.08%) increase in the exchange (network) conversions. This shows that decisions for both channels should be made in tandem after incorporating their interdependencies in the optimization model, and that a failure to do so will result in suboptimal decisions.

⟨⟨Table 4.2 about here⟩⟩

4.4 Structural Estimations

[9] note that if the researcher has sufficient insights into the specific structure of the system, an explicit structural model can be estimated. Along these lines, with the premise of profit maximization by the agency as given in (PM), we develop a structural model, which is based on a system of simultaneous equations involving performance and demand equations.

4.4.1 Assuming Naive Agency

Based on our observations of the decentralized decision making process in both the network and the exchange, as well as anecdotal evidence from the managers in the ad agency, it became evident that there is no recognition of potential complementarities across the channels. Thus, the first step of the estimation involves a structural estimation in which we assume that the agency does not act upon the synergy effects in its profit maximization problem. More specifically, we estimate the following set of equations involving cross-equation restrictions:

$$
\begin{aligned}
y_{E,t} &= A_E^M x_{E,t}^{\alpha_E^M} e^{\kappa_E^M * Trend} e^{u_E^M} \\
y_{N,t} &= A_N^M x_{N,t}^{\alpha_N^M} e^{\kappa_N^M * Trend} e^{u_N^M} \\
x_{E,t}^M &= \left(A_E^M \alpha_E^M \left(\frac{p_{E,t}}{c_{E,t}} \right) e^{\kappa_E^M * Trend} e^{\frac{\left(\sigma_E^M\right)^2}{2}} \right)^{\frac{1}{\left(1-\alpha_E^M\right)}} \\
x_{N,t}^M &= \left(A_N^M \alpha_N^M \left(\frac{p_{N,t}}{c_{N,t}} \right) e^{\kappa_N^M * Trend} e^{\frac{\left(\sigma_N^M\right)^2}{2}} \right)^{\frac{1}{\left(1-\alpha_N^M\right)}}
\end{aligned}
$$

Rewriting the above expressions and taking logarithmic transformations yield a set of linear equations. We also add disturbance terms to the third and fourth equations to explain deviations from the optimizing conditions due to managerial errors. We can estimate the above set of equations through generalized least squares or seemingly unrelated regression (SUR, [59]) where we allow the disturbances to be contemporaeously correlated, i.e., we specify an unrestricted variance-covariance matrix for the error terms. From the above estimation, we compute the fitted values, $\widehat{\ln x_{E,t}}$ and $\widehat{\ln x_{N,t}}$, and we estimate the following two equations

$$
\begin{aligned}
\ln y_{E,t} &= \ln A_E + \alpha_E \widehat{\ln x_{E,t}} + \beta_E \widehat{\ln x_{N,t}} + \kappa_E Trend + v_{E,t} \\
\ln y_{N,t} &= \ln A_N + \alpha_N \widehat{\ln x_{N,t}} + \beta_N \widehat{\ln x_{E,t}} + \kappa_N Trend + v_{N,t}
\end{aligned}
$$

It is worth noting the differences between the coefficients from the first stage estimation, α_E^M and α_N^M, and the coefficients from the second stage estimation, α_E, β_E, α_N, and β_N. The coefficients from the first stage estimation represent the agency's *belief* about the conversion process. As noted above, the agency fails to recognize potential synergy effects between the two channels, and thus believes that the conversion process in each channel is a solely a function of the respective impressions in each channel. On the other hand, the second stage estimation yields the set of coefficients which are of interest to the researcher based on the fitted values from the first stage estimation, which are now free from endogeneity issues.

4.4.2 Assuming Informed Agency

We then move on to empirically quantifying the bias in the coefficients had the researcher misspecified the model by assuming the agency does act upon complementarities. To verify this, we estimate the following set of equations:

$$y_{E,t} = A_E^U x_{E,t}^{\alpha_E^U} x_{N,t}^{\beta_E^U} e^{\kappa_E^U * Trend} e^{u_E^U} \tag{4.11}$$

$$y_{N,t} = A_N^U x_{N,t}^{\alpha_N^U} x_{E,t}^{\beta_N^U} e^{\kappa_N^U * Trend} e^{u_N^U} \tag{4.12}$$

$$\frac{\partial \pi_A}{\partial \pi_E} = p_{E,t} \frac{\partial \mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right]}{\partial x_{E,t}} - c_{E,t} + p_{N,t} \frac{\partial \mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right]}{\partial x_{E,t}} = 0 \tag{4.13}$$

$$\frac{\partial \pi_A}{\partial \pi_N} = p_{E,t} \frac{\partial \mathbb{E}\left[y_{E,t}|x_{E,t}, x_{N,t}\right]}{\partial x_{N,t}} + p_{N,t} \frac{\partial \mathbb{E}\left[y_{N,t}|x_{N,t}, x_{E,t}\right]}{\partial x_{N,t}} - c_{N,t} = 0 \tag{4.14}$$

Notice that the third and fourth equations which represent the first-order conditions for the impression variables are nonlinear. Thus, we utilize an iterative nonlinear SUR estimation technique which we describe below in detail:

Step 1: Using the 2SLS estimates from Section 4.3, A_E, a_E, b_E, A_N, a_N, and b_N, initialize the expected values of y_E and y_N and plug these into Equations (4.13) and (4.14).

Step 2: Estimate the system of equations (4.11) through (4.14) and compute the fitted values $\widehat{y_E}$ and $\widehat{y_N}$.

Step 3: Plug the fitted values, $\widehat{y_E}$ and $\widehat{y_N}$, from Step 2 into Equations (4.13) and (4.14) and re-estimate the system of equations (4.11) through (4.14).

Step 4: Repeat Steps 1 to 3 until the difference between the fitted values from successive iterations is within the prespecified tolerance level.

For estimation in each iteration, we minimize $r' \left(diag(S)_{OLS}^{-1} \otimes I \right) r$ where r is a column vector for the residuals for each equation, S is the variance-covariance matrix[6]. We then compute the fitted values from the iterative nonlinear SUR estimation, $\widehat{\ln x_{E,t}}$ and $\widehat{\ln x_{N,t}}$, and estimate the following two equations

$$\ln y_{E,t} = \ln A_E + \alpha_E \widehat{\ln x_{E,t}} + \beta_E \widehat{\ln x_{N,t}} + \kappa_E Trend + v_{E,t}$$
$$\ln y_{N,t} = \ln A_N + \alpha_N \widehat{\ln x_{N,t}} + \beta_N \widehat{\ln x_{E,t}} + \kappa_E Trend + v_{N,t}$$

Again, there is an important difference between the coefficients from the first estimation, α_E^U, β_E^U, α_N^U, and β_N^U, and those from the second, α_E, β_E, α_N, and β_N. The first set of coefficients refer to the researcher's belief about the agency's understanding of the conversion process, while the second set of coefficients, which is based on the endogeneity-free fitted values from the first-stage estimation, pertain to the actual conversion process. Table 4.3 summarizes the results from the explicit structural estimations. Notice that

[6]The SUR uses the variance-covariance matrix from an OLS solution.

the magnitude of the cross-channel spillover coefficients reduced drastically relative to the IV estimates. For example, the spillover coefficient from the network impressions to the exchange conversions went from 0.1048 in the IV results to 0.0478 and 0.0972 in the explicit structural estimation results. The findings indicate that if the agency ignores synergy effects between the two channels, the resulting decisions are suboptimal, and that the channels may appear to be less complementary regardless of the researcher's assumptions.

⟨⟨Table 4.3 about here⟩⟩

4.5 Verification through Simulations

The results from Section 4.4 are further explored through simulation and counterfactual analysis. Our objective in this section is to investigate how a decision maker's understanding of complementarities and the researcher's knowledge of such understanding can affect the empirical evidence of complementarities or substitutabilities.

4.5.1 Naive Agency

In the first set of simulations, we simulate the agency's decision on the premise that the agency makes independent decisions for each channel. That is, the agency does not take into account the potential cross-channel effects when optimizing for the number of impressions in each channel. The simulation procedure is outlined. We first simulate the agency's decisions using the coefficients, $A_E^M = 1$, $\alpha_E^M = 0.55$, $A_N^M = 1$, and $\alpha_N^M = 0.65$. Specifically, on any given day, the agency's decisions on the number of impressions are generated using Equations (4.7) and (4.8) based on A_E^M, α_E^M, A_N^M, α_N^M, and added variances given the PPA and CPM from each channel. For simplicity, we assume that the randomness associated with deviating from the optimal impression levels are identical and independent. The coefficients used here are the agency's *beliefs* about the conversion processes. Based on the realized numbers of impressions, the number of conversions are simulated using Equations (4.1) and (4.2) with the coefficients, $A_E = 1$, $\alpha_E = 0.5$, $\beta_E = 0.05$, $A_N = 1$, $\alpha_N = 0.6$, and $\beta_N = 0.05$; these coefficients represent *nature* or the true underlying conversion process. On the simulated data set, we then carry out a 2SLS estimation, a structural estimation assuming a naive agency, and another structural estimation assuming an informed agency. We repeat this procedure for a thousand iterations.

Table 4.4 summarizes the simulation results. The simulated data for this table assumes that the agency acts naively, i.e., the agency does not coordinate its decisions across channels. We find that IV estimation was able to recover the coefficients used to simulate the conversions ($\beta_E^{IV} = 0.0510$, $\beta_N^{IV} = 0.495$). On the other hand, the explicit structural estimations failed to do so regardless of the researcher's belief about the agency's behavior. We observe, however, that if the researcher is informed about the agency's behavior, there is less bias in the coefficients ($\beta_E = 0.0450$, $\beta_N = 0.447$) than that of the uninformed researcher who assumes informed agency when in fact the agency is naive ($\beta_E = 0.0602$, $\beta_N = 0.591$). Note also that the direction of the bias is different. We further perform the z-test[7] ([16]) to confirm that indeed both sets of structural estimates are statistically different from the IV estimates.

⟨⟨Table 4.4 about here⟩⟩

4.5.2 Informed Agency

In the second simulation, we generate the agency's decisions assuming that the agency recognizes the potential synergy effects across channels and that it incorporates this into its decision making, i.e., the agency makes coordinated decision making. More precisely, in each iteration the agency's daily decisions are generated using Equations (4.3) and (4.4) with the coefficients, $A_E = 1$, $\alpha_E = 0.5$, $\beta_E = 0.05$, $A_E = 1$, $\alpha_N = 0.6$, and $\beta_N = 0.05$, given the PPA and CPM from each channel. Then, the conversions are generated again using Equations (4.1) and (4.2) with the same set of coefficients. Lastly, we carry out a 2SLS estimation and a structural estimation assuming an informed agency. Table 4.5 shows that under optimal decisions both the 2SLS and the explicit structural model are consistent. Both methods were able to recover the coefficients used to generate the conversions. z-tests on the coefficients confirm that the structural estimates are statistically similar to the IV estimates. The results from Tables 4 and 5 together imply that for the researcher to correctly identify the effect of complementarities through explicit structural estimations, two requirements have to be satisfied concurrently: (1) the de-

[7] The z-test for the difference between two regression coefficients is $z = \frac{\beta_1 - \beta_2}{\sqrt{(SE\beta_1)^2 + (SE\beta_2)^2}}$.

cision maker has to recognize the effect and incorporate it into its decision making, and (2) the researcher has to have a deep understanding of the decision maker's behavior (perhaps from actual observations, as in our study) and apply that knowledge in the structure of the econometric model.

⟨⟨Table 4.5 about here⟩⟩

4.5.3 Ill-informed Agency

Our last set of simulation involves an agency that believes that the two channels, the exchange and network, are substitutes when they are in fact complements. Specifically, in each iteration the impressions for the two channels are generated using Equations (4.3)) and (4.4) with the coefficients, $A_E = 1$, $\alpha_E = 0.6$, $\beta_E = -0.05$, $A_E = 1$, $\alpha_N = 0.7$, and $\beta_N = -0.05$. Then, based on the realized impressions, the conversions are simulated using Equations (4.1) and (4.2) with the coefficients, $A_E = 1$, $\alpha_E = 0.5$, $\beta_E = 0.05$, $A_N = 1$, $\alpha_N = 0.6$, and $\beta_N = 0.05$. On this data set, we run an IV regression and a structural estimation outlined in Section 4.4.2. Table 4.6 summarizes the estimation results. We again find that the IV estimates recover the coefficients, but the structural estimates have a slight upward bias because of which they are statistically different from the IV estimates.

⟨⟨Table 4.6 about here⟩⟩

4.6 Conclusion and Future Research

We analyzed and quantified synergies between channels in an online advertising supply chain in order to study how a decision maker's understanding of complementarity affects the ability of the researcher to extract evidence of such synergies from a structural model. To the best of our knowledge, our study takes the first step of distinguishing between the presence of complementarity and acting upon it to achieve superior performance. We empirically verified that the decision maker's lack of understanding of complementarity can make true complements appear to be substitutes in a structural model. Thus, the utilization of explicit structural models requires a detailed understanding of the decision making process, which can arise out of actually observing how decisions are made in the real world. We find that regardless of the researcher's beliefs about the nature of decision making, complementarities, when present, can be accurately measured in a structural model only when decision makers act upon them. Therefore, researchers should be cautious about the utilization of explicit structural models to measure complementarities. We also showed that IV models will lead to more reliable results, regardless of whether the decision maker is naive or informed about complementarities. Taken together, these results suggest when the researcher is unsure about the decision making process or is aware that potential complementarities may be ignored by the decision maker, the IV approach is likely to produce more accurate results than a structural model.

A key practical implication of our results involves the need to train decision makers to recognize and act upon complementarity (when present), and to even consider reorganizing the decision making structure. For example, we conjecture that the lack of understanding of complementarity across the network and the exchange can partly be attributed to the decentralized nature of decision making by media platform specialists in each channel.

Our future research will focus on developing a theoretical understanding of the bias in the interaction coefficients when the decision maker ignores complementarities. With such a theoretical foundation, we expect to be able to further generalize the result, and to apply our findings to a broad range of problem settings.

4.7 Tables

	Total	Exchange Total	Network Total	Exchange Average	Network Average
Mille Impressions	31,769	22,082	9,687	127	56
Conversions	7,066	4,499	2,567	26	15
Revenue	36,174	23,159	13,015	133	75
Cost	22,707	14,244	8,463	82	49
ROI (%)	59	63	54		
CR (%)	0.0222	0.0204	0.0265		
Pay Per Action				5.14	5.15
Cost Per Mille				0.79	0.98

Table 4.1: Descriptive Statistics

Variables	$\log y_{E,t}$	$\log y_{N,t}$
Intercept	-0.8602***	-1.1682***
$\log x_{E,t}$	0.7828***	0.1074***
$\log x_{N,t}$	0.1048***	0.8327***
Adjusted R^2	0.8725	0.8438

$^{\cdot}p < 0.01$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 4.2: IV Estimation Results

	Variables	$\log y_{E,t}$	$\log y_{N,t}$
Naive Agency	$\log x_{E,t}$	0.8077***	0.0509***
	$\log x_{N,t}$	0.0478$^{\cdot}$	0.8434***
	Adjusted R^2	0.6665	0.8207
Informed Agency (First Stage)	$\log x_{E,t}$	0.7827***	-0.0227*
	$\log x_{N,t}$	0.0187$^{\cdot}$	0.8420***
Informed Agency (Second Stage)	$\log x_{E,t}$	0.7649***	0.1359*
	$\log x_{N,t}$	0.0972$^{\cdot}$	0.8037***
	Adjusted R^2	0.6469	0.8117

$^{\cdot}p < 0.01$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$

Table 4.3: Structural Estimation Results

	Variables	$\log y_{E,t}$	$\log y_{N,t}$
IV Estimation	$\log x_{E,t}$	0.5005	0.0495
	$\log x_{N,t}$	0.0510	0.6016
Naive Agency	$\log x_{E,t}$	0.5489	-
(First Stage)	$\log x_{N,t}$	-	0.6504
Naive Agency	$\log x_{E,t}$	0.5118	0.0447
(Second Stage)	$\log x_{N,t}$	0.0450	0.6222
Informed Agency	$\log x_{E,t}$	0.5516	-0.0015
(First Stage)	$\log x_{N,t}$	-0.0118	0.6574
Informed Agency	$\log x_{E,t}$	0.5093	0.0591
(Second Stage)	$\log x_{N,t}$	0.0602	0.6148

Table 4.4: Structural Estimation Results on Simulated Data Assuming Naive Agency

	Variables	$\log y_{E,t}$	$\log y_{N,t}$
IV Estimation	$\log x_{E,t}$	0.4981	0.0494
	$\log x_{N,t}$	0.0525	0.6025
Informed Agency	$\log x_{E,t}$	0.5182	0.0363
(First Stage)	$\log x_{N,t}$	0.0426	0.6137
Informed Agency	$\log x_{E,t}$	0.4964	0.0520
(Second Stage)	$\log x_{N,t}$	0.0524	0.6092

Table 4.5: Structural Estimation Results on Simulated Data Assuming Informed Agency

	Variables	$\log y_{E,t}$	$\log y_{N,t}$
IV Estimation	$\log x_{E,t}$	0.5000	0.0487
	$\log x_{N,t}$	0.0501	0.5991
Structural Estimation	$\log x_{E,t}$	0.5970	-0.0647
(First Stage)	$\log x_{N,t}$	-0.0415	0.6422
Structural Estimation	$\log x_{E,t}$	0.5241	0.0695
(Second Stage)	$\log x_{N,t}$	0.0588	0.6775

Table 4.6: Structural Estimation Results on Simulated Data Assuming Ill-informed Agency

Appendix

Proofs to Mathematical Results in Chapter 2

Proposition 1. *Assume that* $\bar{\beta}_E$ *and* $\bar{\beta}_N$ *are non-negative, i.e., there are non-negative spillovers across channels. Then the following statements hold.*

1. *The agency's expected profit function* π_A *is supermodular in* $\bar{x}_{E,t}$ *and* $\bar{x}_{N,t}$*.*
2. *The agency's optimal decisions,* $\bar{x}^*_{E,t}$ *and* $\bar{x}^*_{N,t}$*, can be selected such that they are non-decreasing in* $p_{E,t}$ *and* $p_{N,t}$*, and are non-increasing in* $c_{E,t}$ *and* $c_{N,t}$*.*

Proof. 1. Since the expected profit function is twice-continuously differentiable, it suffices to show that the cross-partial of π_A with respect to $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$ is non-negative. Note that

$$\begin{aligned}
\frac{\partial \pi_A}{\partial \bar{x}_{E,t}} &= p_{E,t}\frac{\partial\, \mathbb{E}\left[y_{E,t}\right]}{\partial \bar{x}_{E,t}} + p_{N,t}\frac{\partial\, \mathbb{E}\left[y_{N,t}\right]}{\partial \bar{x}_{E,t}} - c_{E,t} \\
\frac{\partial \pi_A}{\partial \bar{x}_{N,t}} &= p_{E,t}\frac{\partial\, \mathbb{E}\left[y_{E,t}\right]}{\partial \bar{x}_{N,t}} + p_{N,t}\frac{\partial\, \mathbb{E}\left[y_{N,t}\right]}{\partial \bar{x}_{N,t}} - c_{N,t} \\
\frac{\partial^2 \pi_A}{\partial \bar{x}_{E,t}\partial \bar{x}_{N,t}} &= p_{E,t}\frac{\partial^2\, \mathbb{E}\left[y_{E,t}\right]}{\partial \bar{x}_{E,t}\partial \bar{x}_{N,t}} + p_{N,t}\frac{\partial^2\, \mathbb{E}\left[y_{N,t}\right]}{\partial \bar{x}_{E,t}\partial \bar{x}_{N,t}} \geq 0
\end{aligned}$$

Since it is assumed that $\bar{\beta}_E$ and $\bar{\beta}_N$ are non-negative, the above expression is non-negative, and thus π_A is supermodular in $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$.

2. Also, we have

$$\frac{\partial^2 \pi_A}{\partial \bar{x}_{E,t}\partial p_{E,t}} > 0, \frac{\partial^2 \pi_A}{\partial \bar{x}_{E,t}\partial c_{E,t}} < 0, \frac{\partial^2 \pi_A}{\partial \bar{x}_{E,t}\partial p_{N,t}} \geq 0, \frac{\partial^2 \pi_A}{\partial \bar{x}_{E,t}\partial c_{N,t}} = 0,$$

and

$$\frac{\partial^2 \pi_A}{\partial \bar{x}_{N,t}\partial p_{N,t}} > 0, \frac{\partial^2 \pi_A}{\partial \bar{x}_{N,t}\partial c_{N,t}} < 0, \frac{\partial^2 \pi_A}{\partial \bar{x}_{N,t}\partial p_{E,t}} \geq 0, \frac{\partial^2 \pi_A}{\partial \bar{x}_{N,t}\partial c_{E,t}} = 0,$$

Hence, by Topkis's Monotonicity Theorem, the greatest and least elements in the solution pair, $\bar{x}^*_{E,t}$ and $\bar{x}^*_{N,t}$, are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$ and $c_{N,t}$. Moreover, if the solution pair is unique, then the solution pair is guaranteed to satisfy the above properties.

□

Corollary 1.1. *Assume that* $\bar{\beta}_E$ *and* $\bar{\beta}_N$ *are non-positive, i.e., there are non-positive spillovers across channels. Then the following statements hold.*

1. *The agency's expected profit function π_A is submodular in $\bar{x}_{E,t}$ and $\bar{x}_{N,t}$.*
2. *The agency's optimal decision, $\bar{x}^*_{E,t}$ $\left(\bar{x}^*_{N,t}\right)$, can be selected such that it is non-decreasing (non-increasing) in $p_{E,t}$ and $c_{N,t}$, and is non-increasing (non-decreasing) in $p_{N,t}$ and $c_{E,t}$.*

Proposition 2. *Assume that β_E and β_N are non-negative, i.e., there are non-negative spillovers across channels, and that*

$$p_{E,t}\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + (p_{N,t} - w_{N,t})\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} - \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} \geq 0. \qquad \text{(A1)}$$

Then the agency's expected profit function π_A is supermodular in $x_{E,t}$ and $w_{N,t}$, and the agency's optimal decisions, $x^*_{E,t}$ and $w^*_{N,t}$, can be selected such that they are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$. Further, if $w_{N,t} \leq p_{N,t}$, then $x^*_{E,t}$ is non-increasing in $c_{N,t}$.

Proof. Note that

$$\begin{aligned}
\frac{\partial \pi_A}{\partial x_{E,t}} &= p_{E,t}\frac{\partial \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}} + (p_{N,t} - w_{N,t})\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} - c_{E,t} \\
\frac{\partial \pi_A}{\partial w_{N,t}} &= p_{E,t}\frac{\partial \mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + (p_{N,t} - w_{N,t})\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} - \mathbb{E}\left[y_{N,t}\right] \\
\frac{\partial^2 \pi_A}{\partial x_{E,t}\partial w_{N,t}} &= p_{E,t}\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + (p_{N,t} - w_{N,t})\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} - \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}}
\end{aligned}$$

So if the above expression for the cross-partial is non-negative, π_A is supermodular in $x_{E,t}$ and $w_{N,t}$. It also follows that

$$\frac{\partial^2 \pi_A}{\partial x_{E,t}\partial p_{E,t}} > 0, \frac{\partial^2 \pi_A}{\partial x_{E,t}\partial c_{E,t}} < 0, \frac{\partial^2 \pi_A}{\partial x_{E,t}\partial p_{N,t}} \geq 0,$$

and if $w_{N,t} \leq p_{N,t}$ we have

$$\frac{\partial^2 \pi_A}{\partial x_{E,t}\partial c_{N,t}} = p_{E,t}\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial c_{N,t}} + (p_{N,t} - w_{N,t})\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial c_{N,t}} \leq 0$$

We have

$$\frac{\partial^2 \pi_A}{\partial w_{N,t}\partial p_{N,t}} > 0, \frac{\partial^2 \pi_A}{\partial w_{N,t}\partial p_{E,t}} \geq 0, \frac{\partial^2 \pi_A}{\partial w_{N,t}\partial c_{E,t}} = 0.$$

Hence, by Topkis's Monotonicity Theorem, the greatest and least elements in the solution pair, $x^*_{E,t}$ and $w^*_{N,t}$, are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$. If $w_{N,t} \le p_{N,t}$, it also follows that $x^*_{E,t}$ is non-increasing in $c_{N,t}$. Moreover, if the solution pair is unique, then the solution pair is guaranteed to satisfy the above properties. □

Corollary 2.1. *Assume that β_E and β_N are non-positive, i.e., there are non-positive spillovers across channels, and that*

$$p_{E,t}\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \left(p_{N,t} - w_{N,t}\right)\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} - \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} \le 0.$$

Then the agency's expected profit function π_A is submodular in $x_{E,t}$ and $w_{N,t}$, and the agency's optimal decisions, $x^*_{E,t}$ $\left(w^*_{N,t}\right)$, can be selected such that it is non-decreasing in $p_{E,t}$ $(p_{N,t})$, and is non-increasing in $p_{N,t}$ $(p_{E,t})$ and $c_{E,t}$. Further, if $w_{N,t} \le p_{N,t}$, then $x^*_{E,t}$ is non-decreasing in $c_{N,t}$.

Proposition 3. *Assume that β_E and β_N are non-negative, i.e., there are non-negative spillovers across channels, and that*

$$\begin{aligned} p_{E,t}&\left(\frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial^2 \mathbb{E}\left[y_{E,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial \mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right) \\ &- \left(\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} + \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right) + \left(p_{N,t} - w_{N,t}\right) \\ &\left(\frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial^2 \mathbb{E}\left[y_{N,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right) \ge 0. \end{aligned} \quad \text{(A2)}$$

Then the agency's expected profit function π_A is supermodular in $x_{E,t}$ and $w_{N,t}$, and the agency's optimal decisions, $x^{IS}_{E,t}$ and $w^{IS}_{N,t}$, can be selected such that they are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$. Further, if $w_{N,t} \le p_{N,t}$, then $x^{IS}_{E,t}$ is non-increasing in $c_{N,t}$.

Proof. We have

$$
\begin{aligned}
\frac{\partial \pi_A}{\partial x_{E,t}} &= p_{E,t}\left(\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}} + \frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right) \\
&\quad + (p_{N,t} - w_{N,t})\left(\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} + \frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right) - c_{E,t} \\
\frac{\partial \pi_A}{\partial x_{N,t}} &= p_{E,t}\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial w_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + (p_{N,t} - w_{N,t})\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} - \mathbb{E}\left[y_{N,t}\right] \\
\frac{\partial^2 \pi_A}{\partial x_{E,t}\partial w_{N,t}} &= p_{E,t}\left(\frac{\partial^2\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial^2\,\mathbb{E}\left[y_{E,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}\right) \\
&\quad + p_{E,t}\left(\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right) \\
&\quad + (p_{N,t} - w_{N,t})\left(\frac{\partial^2\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}} + \frac{\partial^2\,\mathbb{E}\left[y_{N,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}\right) \\
&\quad + (p_{N,t} - w_{N,t})\left(\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right) \\
&\quad - \left(\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} + \frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right)
\end{aligned}
$$

If the above cross-partial is non-negative, then π_A is supermodular in $x_{E,t}$ and $w_{N,t}$. Moreover, we have

$$\frac{\partial^2 \pi_A}{\partial x_{E,t}\partial p_{E,t}} > 0, \frac{\partial^2 \pi_A}{\partial x_{E,t}\partial c_{E,t}} < 0, \frac{\partial^2 \pi_A}{\partial x_{E,t}\partial p_{N,t}} \geq 0,$$

and if $w_{N,t} \leq p_{N,t}$ we have

$$\frac{\partial^2 \pi_A}{\partial x_{E,t}\partial c_{N,t}} \leq 0$$

We also have

$$\frac{\partial^2 \pi_A}{\partial w_{N,t}\partial p_{N,t}} > 0, \frac{\partial^2 \pi_A}{\partial w_{N,t}\partial p_{E,t}} \geq 0, \frac{\partial^2 \pi_A}{\partial w_{N,t}\partial c_{E,t}} = 0.$$

Hence, by Topkis's Monotonicity Theorem, the greatest and least elements in the solution pair, $x^{IS}_{E,t}$ and $w^{IS}_{N,t}$, are non-decreasing in $p_{E,t}$ and $p_{N,t}$, and are non-increasing in $c_{E,t}$. If $w_{N,t} \leq p_{E,t}$, it also follows that $x^{IS}_{E,t}$ is non-increasing in $c_{N,t}$. Moreover, if the solution pair is unique, then the solution pair is guaranteed to satisfy the above properties. □

Corollary 3.1. *Assume that β_E and β_N are non-positive, i.e., there are non-positive spillovers across channels, and that*

$$p_{E,t}\left(\frac{\partial^2\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+\frac{\partial^2\,\mathbb{E}\left[y_{E,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+\frac{\partial\,\mathbb{E}\left[y_{E,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right)$$
$$-\left(\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}}+\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right)$$
$$+\left(p_{N,t}-w_{N,t}\right)\left(\frac{\partial^2\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}+\frac{\partial^2\,\mathbb{E}\left[y_{N,t}\right]}{\partial x^2_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}\right)$$
$$+\left(p_{N,t}-w_{N,t}\right)\left(\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial^2 x^*_{N,t}}{\partial x_{E,t}\partial w_{N,t}}\right)\leq 0.$$

Then the agency's expected profit function π_A is submodular in $x_{E,t}$ and $w_{N,t}$, and the agency's optimal decisions, $x^{IS}_{E,t}$ $\left(w^{IS}_{N,t}\right)$, can be selected such that it is non-decreasing in $p_{E,t}$ $(p_{N,t})$, and is non-increasing in $p_{N,t}$ $(p_{E,t})$ and $c_{E,t}$. Further, if $w_{N,t}\leq p_{N,t}$, then $x^{IS}_{E,t}$ is non-increasing in $c_{N,t}$.

Proposition 4. *Let $k=\frac{c_{N,t}x^*_{N,t}}{w_{N,t}\,\mathbb{E}\left[y_{N,t}\right]}$, the publisher's cost-benefit ratio (cost over expected revenue). If $k\leq 1$, then the optimal decisions, x^{PPS}_E and w^{PPS}_N, can be selected such that it is non-decreasing in γ.*

Proof. We have

$$\frac{\partial^2\pi_A}{\partial x_{E,t}\partial\phi}=-\left(w_{N,t}\left(\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}}+\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right)-c_{N,t}\frac{\partial x^*_{N,t}}{\partial x_{E,t}}\right)$$
$$\frac{\partial^2\pi_A}{\partial w_{N,t}\partial\phi}=-\left(\mathbb{E}\left[y_{N,t}\right]+w_{N,t}\frac{\partial\,\mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}-c_{N,t}\frac{\partial x^*_{N,t}}{\partial w_{N,t}}\right)$$

Rearranging the terms we conclude that by Topkis's Monotonicity Theorem the greatest and least elements in the solution pair, $x^{PPS}_{E,t}$ and $w^{PPS}_{N,t}$, are non-decreasing in ϕ if

$$k\equiv\frac{c_{N,t}x^*_{N,t}}{w_{N,t}\,\mathbb{E}\left[y_{N,t}\right]}\leq 1.$$

Moreover, if the solution pair is unique, then the solution pair is guaranteed to satisfy the above property. □

Proposition 5. *Let $k'=\frac{c_{N,t}x^*_{N,t}}{p_{N,t}\,\mathbb{E}\left[y_{N,t}\right]}$, the cost-benefit ratio (cost over expected revenue). If $k'\leq 1$, then the optimal decision, x^{APS}_E, can be selected such that it is non-decreasing in γ.*

Proof. Note that

$$\frac{\partial^2 \pi_A}{\partial x_{E,t} \partial \phi} = p_{N,t} \left(\frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{E,t}} + \frac{\partial \mathbb{E}\left[y_{N,t}\right]}{\partial x_{N,t}} \frac{x^*_{N,t}}{x_{E,t}} \right) - c_{N,t} \frac{\partial x^*_{N,t}}{\partial x_{E,t}}$$

Rearranging the terms we conclude that by Topkis's Monotonicity Theorem the greatest and least elements in the solution, $x^{APS}_{E,t}$, are non-decreasing in ϕ if

$$k^{'} \equiv \frac{c_{N,t} x^*_{N,t}}{p_{N,t}\, \mathbb{E}\left[y_{N,t}\right]} \leq 1.$$

Moreover, if the solution pair is unique, then the solution pair is guaranteed to satisfy the above property. $\square$

Bibliography

[1] Vibhanshu Abhishek, Peter Fader, and Kartik Hosanagar. Media exposure through the funnel: A model of multi-stage attribution. *Working Paper*, 2012.

[2] Joshua Angrist and Alan B Krueger. Why do world war ii veterans earn more than nonveterans? *Journal of labor economics*, 12(1):74–97, 1994.

[3] Ashish Arora. Testing for complementarities in reduced-form regressions: A note. *Economics Letters*, 50(1):51–55, 1996.

[4] Susan Athey and Scott Stern. An empirical framework for testing theories about complimentarity in organizational design. *Working Paper*, 1998.

[5] Robert L Bray and Haim Mendelson. Information transmission and the bullwhip effect: An empirical investigation. *Management Science*, 58(5):860–875, 2012.

[6] Trevor S Breusch. Maximum likelihood estimation of random effects models. *Journal of Econometrics*, 36(3):383–389, 1987.

[7] Carol V Brown and Sharon L Magill. Reconceptualizing the context-design issue for the information systems function. *Organization Science*, 9(2):176–194, 1998.

[8] Erik Brynjolfsson and Lorin Hitt. Paradox lost? firm-level evidence on the returns to information systems spending. *Management Science*, 42(4):541–558, 1996.

[9] Erik Brynjolfsson and Paul Milgrom. Complementarity in organizations. *The handbook of organizational economics*, pages 11–55, 2013.

[10] Gérard P Cachon. *Supply chain coordination with contracts*, volume 11. Elsevier, 2003.

[11] Gérard P Cachon and Marshall Fisher. Supply chain inventory management and the value of shared information. *Management Science*, 46(8):1032–1048, 2000.

[12] Gérard P Cachon and Martin A Lariviere. Supply chain coordination with revenue-sharing contracts: strengths and limitations. *Management Science*, 51(1):30–44, 2005.

[13] Donald T Campbell and Julian C Stanley. *Experimental and quasi-experimental designs for research*. Ravenio Books, 2015.

[14] Fangruo Chen. Echelon reorder points, installation reorder points, and the value of centralized demand information. *Management Science*, 44(12-part-2):S221–S234, 1998.

[15] Fangruo Chen. *Information sharing and supply chain coordination*, volume 11. Elsevier, 2003.

[16] Clifford C Clogg, Eva Petkova, and Adamantios Haritou. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology*, 100(5):1261–1293, 1995.

[17] Thomas D Cook, Donald Thomas Campbell, and William Shadish. *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin Boston, 2002.

[18] Ruomeng Cui, Gad Allon, Achal Bassamboo, and Jan A Van Mieghem. Information sharing in supply chains: An empirical and theoretical valuation. *Management Science*, 61(11):2803–2824, 2015.

[19] Ming Fan, Jan Stallaert, and Andrew B Whinston. Decentralized mechanism design for supply chain organizations using an auction market. *Information Systems Research*, 14(1):1–22, 2003.

[20] Anindya Ghose and Sha Yang. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, 55(10):1605–1622, 2009.

[21] Avi Goldfarb and Catherine Tucker. Advertising bans and the substitutability of online and offline advertising. *Journal of Marketing Research*, 48(2):207–227, 2011.

[22] Avi Goldfarb and Catherine Tucker. Online display advertising: Targeting and obtrusiveness. *Marketing Science*, 30(3):389–404, 2011.

[23] Avi Goldfarb and Catherine Tucker. Search engine advertising: Channel substitution when pricing ads to context. *Management Science*, 57(3):458–470, 2011.

[24] William H Greene. *Econometric analysis*. Pearson Education India, 2003.

[25] Sang-Pil Han, Anindya Ghose, and Sung-Hyuk Park. Cross-channel synergies between web and mobile advertising: A randomized field experiment. *Working Paper*, 2013.

[26] Andrew C Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.

[27] Yu Hu, Jiwoong Shin, and Zhulei Tang. Incentive problems in performance-based online advertising pricing: cost per click vs. cost per action. *Management Science*, 62(7):2022–2038, 2015.

[28] Guido W Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635, 2008.

[29] Pavel Kireyev, Koen Pauwels, and Sunil Gupta. Do display ads influence search? attribution and dynamics in online advertising. *International Journal of Research in Marketing*, 33(3):475–490, 2016.

[30] Martin A Lariviere and Evan L Porteus. Selling to the newsvendor: An analysis of price-only contracts. *Manufacturing & service operations management*, 3(4):293–305, 2001.

[31] Hongshuang Li and PK Kannan. Attributing conversions in a multichannel online marketing environment: An empirical model and a field experiment. *Journal of Marketing Research*, 51(1):40–56, 2014.

[32] Manoj K Malhotra and Alan W Mackelprang. Are internal manufacturing and external supply chain flexibilities complementary capabilities? *Journal of Operations Management*, 30(3):180–200, 2012.

[33] Puneet Manchanda, Jean-Pierre Dubé, Khim Yong Goh, and Pradeep K Chintagunta. The effect of banner advertising on internet purchasing. *Journal of Marketing Research*, 43(1):98–108, 2006.

[34] Larry J Menor, M Murat Kristal, and Eve D Rosenzweig. Examining the influence of operational intellectual capital on capabilities and performance. *Manufacturing & Service Operations Management*, 9(4):559–578, 2007.

[35] Paul Milgrom and John Roberts. The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review*, pages 511–528, 1990.

[36] Suresh Muthulingam and Anupam Agrawal. Does quality knowledge spillover at shared suppliers? an empirical investigation. *Manufacturing & Service Operations Management*, 18(4):525–544, 2016.

[37] Ram Narasimhan, Morgan Swink, and Sridhar Viswanathan. On decisions for integration implementation: An examination of complementarities between product-process technology integration and supply chain integration. *Decision Sciences*, 41(2):355–372, 2010.

[38] Sharon Novak and Steven D Eppinger. Sourcing by design: Product complexity and the supply chain. *Management Science*, 47(1):189–204, 2001.

[39] Sharon Novak and Scott Stern. Complementarity among vertical integration decisions: Evidence from automobile product development. *Management Science*, 55(2):311–332, 2009.

[40] Marcelo Olivares, Christian Terwiesch, and Lydia Cassorla. Structural estimation of the newsvendor model: an application to reserving operating room time. *Management Science*, 54(1):41–55, 2008.

[41] Nikolay Osadchiy, Vishal Gaur, and Sridhar Seshadri. Systematic risk in supply chain networks. *Management Science*, 62(6):1755–1777, 2015.

[42] Barry Alan Pasternack. Optimal pricing and return policies for perishable commodities. *Marketing Science*, 4(2):166–176, 1985.

[43] Arun Rai, Ravi Patnayakuni, and Nainika Seth. Firm performance impacts of digitally enabled supply chain integration capabilities. *MIS quarterly*, pages 225–246, 2006.

[44] Kevin Ross and Kristin Fridgeirsdottir. Cost-per-impression and cost-per-action pricing in display advertising with risk preferences. *Working Paper*.

[45] Oliver J Rutz, Randolph E Bucklin, and Garrett P Sonnier. A latent instrumental variables approach to modeling keyword conversion in paid search advertising. *Journal of Marketing Research*, 49(3):306–319, 2012.

[46] Juan Camilo Serpa and Harish Krishnan. The impact of supply chains on firm-level productivity. *Management Science*, 64(2):511–532, 2017.

[47] Nicolaj Siggelkow and Daniel A Levinthal. Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational approaches to exploration and adaptation. *Organization Science*, 14(6):650–669, 2003.

[48] Nicolaj Siggelkow and Jan W Rivkin. Speed and search: Designing organizations for turbulence and complexity. *Organization Science*, 16(2):101–122, 2005.

[49] Jitendra V Singh. Performance, slack, and risk taking in organizational decision making. *Academy of management Journal*, 29(3):562–585, 1986.

[50] Joseph J Spengler. Vertical integration and antitrust policy. *Journal of political economy*, 58(4):347–352, 1950.

[51] Douglas Staiger and James H Stock. Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557, 1997.

[52] Hüseyin Tanriverdi. Performance effects of information technology synergies in multibusiness firms. *Mis Quarterly*, pages 57–77, 2006.

[53] Donald M Topkis. *Supermodularity and complementarity.* Princeton university press, 2011.

[54] Catherine E Tucker. Social advertising: How advertising that explicitly promotes social influence can backfire. *Working Paper*, 2016.

[55] Yixin (Iris) Wang, Jun Li, and Ravi Anupindi. Risky suppliers or risky supply chains? an empirical analysis of sub-tier supply network structure on firm risk in the high-tech sector. *Working Paper*, 2017.

[56] Elaine M Wong, Margaret E Ormiston, and Philip E Tetlock. The effects of top management team integrative complexity and decentralized decision making on corporate social performance. *Academy of Management Journal*, 54(6):1207–1228, 2011.

[57] Lizhen Xu, Jason A Duan, and Andrew Whinston. Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science*, 60(6):1392–1412, 2014.

[58] Sha Yang and Anindya Ghose. Analyzing the relationship between organic and sponsored search advertising: Positive, negative, or zero interdependence? *Marketing Science*, 29(4):602–623, 2010.

[59] Arnold Zellner. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American statistical Association*, 57(298):348–368, 1962.

[60] Arnold Zellner, Jan Kmenta, and Jacques Dreze. Specification and estimation of cobb-douglas production function models. *Econometrica: Journal of the Econometric Society*, pages 784–795, 1966.

[61] Arnold Zellner and Henri Theil. Three-stage least squares: simultaneous estimation of simultaneous equations. *Econometrica: Journal of the Econometric Society*, pages 54–78, 1962.

[62] Kevin Zhu. The complementarity of information technology infrastructure and e-commerce capability: A resource-based assessment of their business value. *Journal of Management Information Systems*, 21(1):167–202, 2004.

Vita

Changseung Yoo was born in Daejeon, South Korea, the son of Yong Ok Yoo and Ho Sook Kim. He received the Bachelor of Business Administration degree from Yonsei University and the Master of Science degree in Operations Research from Columbia University. He then applied to the University of Texas at Austin for enrollment in their Ph.D. in Information, Risk, and Operations Management Program. He was accepted and started graduate studies in August, 2012.

Email address: csyoo83@gmail.com

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.