

Catching (Exclusive) Eyeballs:
Multi-Homing and Platform Competition in the Magazine Industry

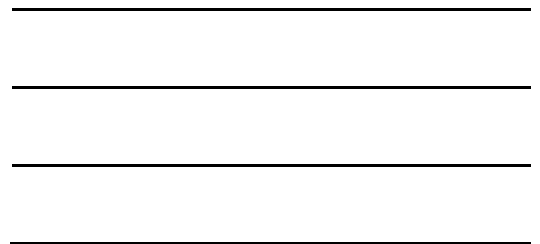
Ce Shi
City of Hangzhou, Zhejiang Province, P.R.China

M.A. in Economics, University of Virginia, 2011
B.S. in Mathematics and Economics, University of Illinois at Urbana-Champaign, 2008

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy for the Degree of
Doctor of Philosophy

Department of Economics

University of Virginia
August, 2016



Catching (Exclusive) Eyeballs: Multi-Homing and Platform Competition in the Magazine Industry

Ce Matthew Shi¹

June 26, 2016

¹Department of Economics, University of Virginia. E-mail: *matthew.shi@virginia.edu*. I am deeply indebted and extremely grateful to Simon Anderson, Steven Stern, and Federico Ciliberto for their continuous guidance and encouragement. I thank my fourth reader, Natasha Zhang Foutz, for her insightful comments. This work has benefited from conversations with Nathan Larson, Regis Renault, Jesse Shapiro, and Michael Zheng Song as well as comments from Oystein Foros, Hans Jarle Kind, Charles Murry, and seminar participants at Bates White, the Chinese University of Hong Kong, University of Virginia, and Zhejiang University. I would also like to thank Summer Durrant (UVa Alderman Library), Angela La Grasta (GfK MRI) and Liz Grabler (MA-Focus Media) for help with data access. I acknowledge financial support from the Bankard Fund for Political Economy. All errors are my own.

Media platforms compete for both consumers and advertisers, especially when consumers divide their attention among multiple platforms. In this dissertation, I study how media competition affects media advertising volume and revenues. In particular, I investigate how consumer multiple purchasing behavior (or “multi-homing”) shapes competition in two-sided media markets. Standard models of two-sided markets have often been based on the assumption that consumers limit their attention to a single platform (or “single-homing”). As applied to media markets, it implies that media platforms have a monopoly position over the consumer impression. While insightful, this approach has been challenged by many empirical puzzles. This dissertation joins a nascent theoretical literature (e.g., Ambrus, Calvano, and, Reisinger (2016); Anderson, Foros, and, Kind (2016)) to incorporate consumer multi-homing into an empirical model of two-sided magazine markets.

Specifically, I build and estimate a structural model of consumer demand for multiple magazines, advertiser demand and magazines two-sided pricing decisions. I compile a novel data-set of detailed MSA-level magazine sales, advertising quantities and prices, and consumers state order of preference of major U.S. magazines in six genres from 2003 to 2012. I use the Method of Simulated Moments (MSM) to estimate the model, quantifying the cross-group externalities in magazine markets. On the reader side, I find that consumer ad nuisance cost is approximately 5 cents per ad page, in contrast to the ad-neutrality or ad-loving findings in the literature on print media. On the advertiser side, I estimate that, on average, advertisers value exclusive eyeballs at 12 cents each, or more than twice the value of a shared consumer. This is the first direct evidence that media ad prices reflect advertisers differential valuation of exclusive and shared consumers on platforms, supporting the hypothesis in the theoretical literature. In a counter-factual exercise, I illustrate how multi-homing consumers affect market outcomes as the market environment for print media has

worsened over the decade. I investigate how the market outcomes would have changed if demand for magazines in 2012 remained as strong as in 2003. I interpret the results as the effects of the Internet on magazine markets. Subscription levels would increase by 9 percent on average despite higher subscription prices. Exclusive readerships would increase by 22 percent, which translates to stronger advertising demand and a 37-percent increase in ad revenues. The internet has large impacts on magazine ad markets by reducing the number of consumers and also by changing the consumer composition on platforms.

JEL CLASSIFICATION: C25, D12, D22, L13, L82, M37

KEYWORDS: Two-sided markets; media economics; advertising; platform competition; magazine industry

Contents

1	Introduction	1
2	Literature Review	6
2.1	Related Literature in Industrial Organization Theory	6
2.2	Related Literature in Empirical Industrial Organization	11
2.3	Related Literature on Modeling Multiple Consumer Choices	16
3	Industry Background	19
4	The Model	24
4.1	Overview of the Model	24
4.2	Demand for Magazines	25
4.2.1	Choice Probability and Market Penetration Function	27
4.2.2	Further Discussion on κ and the $n = 4$ Assumption	30
4.2.3	Regularity Conditions for the Reader Demand	31
4.2.4	Exclusive Readership	33
4.2.5	Discussion on the Modeling Choice	34
4.3	Demand for Advertising	35
4.4	Magazine Platforms	38
5	Data	41
5.1	Data Description	41

	IV
5.2 Descriptive Evidence	44
6 Empirical Implementation	48
6.1 Econometric Specifications	48
6.1.1 Specification of the Reader Demand Model	48
6.1.2 Specification of the Advertiser Demand Model	50
6.2 Estimation of the Empirical Model	51
6.2.1 Moments	52
6.2.2 Identification	55
7 Estimation Results and the Counterfactual Analysis	62
7.1 Estimation Results	62
7.1.1 Reader Demand Estimates	62
7.1.2 Advertiser Demand Estimates	71
7.2 Counterfactual Exercise: The Impact of the Internet	73
8 Conclusion	76
Appendices	84

Chapter 1

Introduction

In the rapidly changing media landscape, a prominent question is how competition affects media advertising volume and revenues. This is especially important as traditional media rely more heavily on advertising revenues to finance their platforms and to provide valuable content to media consumers. In order to understand this, one needs to recognize the “two-sidedness” of media markets. When markets are two-sided, two distinct groups of users interact via platform firms. In the case of media markets, on one side are media consumers, such as magazine and newspaper readers, television viewers and web surfers. On the other side are advertisers who want to get consumer attention for sales of product. In between, media platforms, such as magazines and newspaper, television stations and online platforms, compete to attract consumer attention and in turn charge advertisers for getting into the attention of their media consumers, or metaphorically “eyeballs”. Standard economic analyses of media markets have been based on the assumption that consumers limit their attention to one single platform, which is called “single-homing” (SH) in the literature. As described in detail in the literature review chapter of my thesis, Chapter 2, this single-homing assumption is frequently made for simplicity and tractability of theoretical analysis but has important implications for modeling of advertiser behav-

ior and platform competition as well as for drawing useful policy recommendations. In particular, when media consumers are single-homers, media platforms are essentially monopolists over the fresh impression of their consumers (e.g., Anderson and Coate (2005)). While insightful, this single-homing approach has been challenged by empirical puzzles that the theory cannot explain satisfactorily or that contradict predictions of the theory. As discussed in Chapter 2, a standard media economics model with single-homing consumers predicts that a merger between two media platforms reduces prices of advertising; that entry of an additional platform raises ad prices; and that when a public media provider is allowed to carry ads, the private ones are better off. Moreover, the single-homing assumption simply does not account for the fact that consumers often patronize multiple platforms for content needs (or “multi-home”, MH) in reality. As described in detail in the chapter on industry background of U.S. magazines, Chapter 3, U.S. households commonly subscribe to between 2 and 3 magazines annually. And this rough average number does not account for the facts that conditioned on being a subscriber, the average number is be higher, and that consumers often have other ways to obtain copies of magazines to read - such as those in libraries, medical clinics, and barber shops - besides private subscription. In response to the disagreements between theoretical predictions and empirical observations, new theoretical works emerge to revise the standard ad-financed model of media markets. One relatively small and nascent literature, including Ambrus, Calvano, and Reisinger (2016), Anderson, Foros, and Kind (2016), and Athey, Calvano, and Gans (2014), relaxes the assumption of single-homing consumers. Joining this literature, this dissertation incorporates consumer multi-homing and other insights from the theoretical frontier into an empirical model of two-sided magazine markets. Presented in the main chapters - Chapters 4, 5, 6 and 7 - of this thesis, I estimate the model with novel data on metro-level magazine sales, advertising quantities and prices, and consumers’ stated order of preference to quantify the cross-group external-

ities in magazine markets. I demonstrate the importance of consumer multi-homing in estimating the reader demand and in shaping ad prices of platforms. In a counterfactual exercise, I illustrate how multi-homing consumers affect subscription and advertising market outcomes in a decade of Internet expansion.

In addition to its implications for the theory of two-sided media markets, consumer multi-homing behavior has important implications for empirical modeling of media markets and estimation. First, traditional discrete choice models used in demand estimation very often assume that consumers buy only one product among some set of differentiated products. This single-purchasing assumption in discrete choice models - which corresponds nicely to the single-homing assumption in two-sided market contexts - can be problematic if the true data-generating process is characterized with consumer multiple purchasing activities. In the model chapter, Chapter 4, following Hendel (1999) and Fan (2013), I set up a multiple discrete choice model of magazine readers. I show how a traditional discrete choice model can be modified to handle consumer multi-homing behavior, and I verify that the resulted reader demand functions do satisfy regularity conditions needed for both theoretical consonance and estimation. In the empirical parts of this thesis, I use a new data set that I collect from multiple sources, including detailed magazine sales at the Metropolitan Statistical Area (MSA) level in six major genres for ten years, magazine characteristics, and survey data that reveal consumers' order of preference for magazines that they have purchased. I construct moments from the novel survey data to identify a new parameter that captures consumer utility loss from multiple purchases, which is generally unidentified using only market-level data. In contrast to findings in the established literature on print media, I find that consumers dislike magazine advertisements. In particular, one more page of ad is equivalent to an increase in subscription price by five cents, on average. In addition, I find heterogeneity in consumer ad nuisance and in the decision whether or not to purchase magazines at all. I then compare the

results to those from standard logit and mixed logit models with single-purchasing consumers. The single-homing models significantly underestimate consumer distastes for advertising and overestimates their tastes for other favorable product attributes.

Second, when consumers multi-home in media markets, they spread their attention among multiple platforms. If advertisers' return to consumer attention diminishes after the first impression, their willingness to pay for exclusive consumers and for shared consumers on a platform should differ. Consequentially, platform prices of advertising should reflect this advertisers' differential valuation of consumers, *ceteris paribus*. In fact, these key insights from the somewhat nascent theoretical literature have been long observed by practitioners in advertising and media industries. Regarding this point, Martin (1921, cf. Gentzkow, Shapiro, and Sinkinson 2014) writes that "*The same advertisement seen in two or three newspapers is certainly more effective than if seen in one, but some advertisers are convinced that it is not worth three times as much to have an advertisement seen in three papers, reaching largely the same readers, as to have it seen in one.*" Reflecting that observation, it is not uncommon that media compare their exclusive readerships to their closest substitutes in effort to sell advertising space. For instance, a magazine states upfront to potential advertisers that "*[Their] advertising rates are extremely cost efficient. [Their] targeted and exclusive readership means there is very little wastage.*" (Highlife Magazine 2015).¹ Motivated by these observations, on the advertiser side, I estimate an inverse demand function for advertising, including both the subscription level and the number of exclusive/single-homing consumers as explanatory variables. In particular, I use the reader-side model and the estimates to predict missing data on the number of exclusive consumers on each platform, and use the predicted values to help explain advertiser demand. I find shared consumers worth 4.9 cents per eyeball while exclusive consumers worth 12 cents or twice as much. This is one of the first evidence to

¹See <http://highlifemagazine.com.au/advertise-with-us/>.

support the major hypothesis in the theoretical two-sided market literature. Subsequently, I conduct a counterfactual analysis supposing demand for magazines at the end of my sample period were as strong as in 2003. The results suggest that the Internet has affected the magazine markets not only by lowering revenues from circulation but also by reducing the number of subscribers, and thus making the reader composition less favorable to platforms.

This thesis is organized as follows. In the next chapter, Chapter 2, I review three strands of literature related to the theory of two-sided markets, the empirical studies of media, and empirical modeling of consumer choice with multiple purchases. Chapter 3 summarizes some key features of the magazine industry. In Chapter 4, I present a structural model of reader demand, advertiser demand and magazine decisions. Chapter 5 discusses the data used in this project with descriptive evidence. Chapter 6 explains the estimation procedure and discusses identification issues. Chapter 7 represents results from the estimation and from a counterfactual exercise. Finally, Chapter 8 concludes.

Chapter 2

Literature Review

2.1 Related Literature in Industrial Organization Theory

Despite the economic importance of media markets to consumers, producers, and advertisers - including unusual advertisers like politicians, media industries had not received much special attention in the theoretical Industrial Organization literature before introduction of the modeling framework of two-sided markets. Earlier works on media markets often focus on roles of advertising in the media (e.g., Steiner (1952) and Beebe (1977)). It was when the theoretical literature distinguishes two-sided markets from traditional “one-sided” markets that media markets became a prominent example of two-sided markets. Foundational works on the two-sided market theory in Industrial Organization include Caillaud and Jullien (2003), Rochet and Tirole (2003), and Armstrong (2006). According to Armstrong (2006), two-sided markets are “where two or more groups of agents interact via intermediaries or ‘platforms’.” A distinct feature of two-sided markets is the presence of cross-group externalities. In a typical model, each side’s willingness to join a platform depends on the number of participants on the other side and the price of joining. When a group of agent

joins a single platform, it is called “single-homing”. Analogously, joining multiple platforms is called “multi-homing”. Armstrong (2006) lists these three features - cross-group externalities, platform prices, and single-homing/multi-homing - as the main determinants of the pricing structure in two-sided markets. Besides the traditional media, such as television, radio, newspaper, and magazines, other examples of two-sided markets include credit cards, video game consoles, social media, and shopping malls. Armstrong (2006) surveys three types of market structures that he regards as important: two-sided monopoly, two-sided single-homing, and the “competitive bottlenecks”, in which one side of users single-homes, and the other side multi-homes. In all of these situations, it is usually assumed that platforms compete by choosing the number of participants on each side. In addition, in many cases, it is assumed that both sides exert positive externalities on the other sides. Noticeably, the case of two-sided multi-homing has not been well studied in those early general two-sided market models. As Armstrong (2006) puts, “[if] interacting with the other side is the primary reason for an agent to join a platform, then we might not expect case [of two-sided multi-homing] to be very common: if each member of group 2 joins all platforms, there is no need for any member of group 1 to join more than one platform.” However, it turns out that, in many media markets, not only two-sided multi-homing is prevalent, but also it matters critically for drawing policy predictions.

Anderson and Coate (2005) develops a theoretical model of a two-sided media market, which in many ways has become the standard model for the analysis of media markets. In the context of broadcasting industries, they assume a competitive bottleneck structure, in which consumers single-home and advertisers multi-home. This assumption is closely followed by many subsequent theoretical and structural empirical works. On the audience side, they build a Hotelling-type model, in which two differentiated broadcasting stations or platforms compete for consumers in a “linear city,” and consumers who care about the volume of advertising carried in each sta-

tion choose the station that yields higher utilities. Although introduced originally in a somewhat different theoretical context, the single-homing consumer assumption is especially well-received by structural empiricists because it nicely fits into the empirical discrete choice models with single-purchasing consumers, including Berry, Levinhson, and Pakes (1995). More recent theoretical works, such as Anderson and Peitz (2014) and Anderson et al. (2014), that uses a multinomial logit model on the consumer side also assumes a competitive bottleneck structure. Another influential assumption made in Anderson and Coate (2005) is consumer ad-aversion. Consumer ad-aversion means that media consumers prefer non-ad programs or content to advertising for their leisure time. However, it does not necessarily imply ineffectiveness of advertising messages: average consumers may still find advertising useful in some way, even though they find it less enjoyable than non-ad programs. Furthermore, ad-aversion means that the participation of advertisers would impose a negative externality on consumers - an important feature that distinguishes many media markets from other two-sided markets, where each side generates a positive externality on the other, such as credit cards, night clubs, and yellow pages. Because of its importance, the question whether consumers dislike ads or not has become a heated topic especially in the empirical literature. As discussed in a later section, researchers have found ambiguous results on consumer ad-aversion, basing on empirical investigations of many media forms in different countries. This thesis joins the debate by offering evidence from the U.S. magazine markets that supports the consumer ad-aversion assumption.

On the advertiser side, Anderson and Coate (2005) derives a linear (inverse) ad demand equation from the fundamentals of advertising technology and advertiser payoff function. Although their characterization of advertisers is somewhat stylized, their presentation of advertising demand is nonetheless highly general, applicable to any characterization of advertisers that would lead to a linear ad demand equation. A few features of the advertiser model are worth noting. First, advertisers are assumed

to be monopoly producers of their products. In other words, competition *among* advertisers is suppressed, and there is no strategic behavior of advertisers. Up to the date of this thesis, almost all theoretical and empirical research on two-sided media markets have taken this approach to assume away advertiser competition, given the complexities of the current model and coarseness of advertising data used in empirical studies. However, it is also not difficult to see why advertisers' decisions might be interdependent. For example, it is common in the pharmaceutical industry that drug producers launch ads to attack competing products and to self-promote (Anderson et al. 2016). Allowing for advertising competition in a two-sided media market model is an interesting subject for further study. Second, in conjunction with the first point, because of the single-homing assumption on the consumer side, platforms have an actual monopoly position over the consumer attention to advertisers. The resulting ad demand function is of a monopoly one, which gives much convenience for the empirical analysis based on aggregate advertising price and/or quantity data. Third, by assuming that advertising price is charged at a per-viewer base, it avoids many complexities of solving the model that are found in the general two-sided market framework (Armstrong 2006). While this assumption is appreciated in the theoretical literature (Weyl and White (2014)), it receives less attention in the empirical literature. In particular, when a per-viewer advertising price is assumed, it forced the resulting ad pricing equation to be linear in the size of audience. In fact, as noted by Athey, Calvano, and Resinger (2014), this is not the case in some prominent media markets: anecdotal evidence suggests that larger media platforms tend to charge higher ad prices per-viewer, which phenomenon gives names of term "Fox News Puzzle" or "ITV premium puzzle". While this thesis adopts the same assumption regarding the first point, I take innovation steps to explore the possibility of relaxing the other assumptions in an empirical model of two-sided media.

In Anderson and Jullien (2016)'s recent survey article on the business model of

tow-sided media, they summarize that the single-homing consumer assumption in competitive bottlenecks may give rise to some unusual policy predictions: for example, the competitive bottleneck model with one-sided pricing predicts that mergers would increase advertising volume in media but decrease advertising prices. The intuition is rather simple: no matter how competition in consumer markets is modeled (spatial or discrete choice model), as long as consumers single-home, platforms are monopolists in the advertising markets. As in a traditional one-sided market, mergers increase the market power of platforms, enabling media to make consumers pay higher full price by introducing more ad nuisances. In a monopoly, higher advertising volume translates automatically to lower advertising prices.

As pointed out by Anderson and Jullien (2016), relaxing the consumer single-homing assumption may help improve predictions of the theoretical model.¹ In response to the empirical puzzles mentioned above, there is an emerging theoretical literature that relaxes the single-homing assumption. The key idea of this literature is that when consumers may multi-home, platforms also face competition in the advertising markets since they are no longer the sole conduits for delivering their consumer eyeballs. Furthermore, if there is diminishing return to consumer impression, advertisers must value exclusive consumers - those who are accessible via a single platform - more than overlapping consumers who divide their attention among multiple media outlets. A pioneering theoretical treatment is Ambrus, Calvano, and Reisinger (2015),² where they introduce a duopoly model of platform competition in both the consumer and advertising markets. In a similar spirit, Anderson, Foros, and Kind (2016) develops a model with n platforms competing for advertisers a la Bertrand. Because both sides may multi-home, price competition between plat-

¹Noticeably, other modifications can be made to the standard one-sided pricing, ad-financed model to generate different policy predictions. For example, in a two-sided pricing model with consumer single-homing, such empirical puzzles may not exist (Anderson and Peitz (2015)).

²The published version of Ambrus, Calvano, and Reisinger (2015) replaces Ambrus and Reisinger (2007) - a long-circulated working paper on the same subject.

forms implies that, in equilibrium, a platform’s advertising price should depend on the number of exclusive consumers it delivers to advertisers and not on the number of shared consumers. This is termed the “incremental pricing principle”. Therefore, as two platforms merge, some formerly shared eyeballs now become exclusive to the merged platform, thus giving it higher ability to price beyond marginal cost. Similarly, this model with two-sided multi-homing resolves other puzzles related to market entry/exit and public media carrying ads. This thesis joins this small literature by introducing two-sided multi-homing to an empirical model of media markets and providing the first direct evidence that advertisers value exclusive consumers more than the shared eyeballs.

2.2 Related Literature in Empirical Industrial Organization

In the empirical Industrial Organization literature, Rysman (2004) is the first paper to estimate a structural two-sided market model. Using a cross-sectional dataset on the U.S. yellow page markets in 1996, the author quantifies the cross-group externalities presented in the market. He finds positive externalities exhibited on both sides of the market: while advertisers value more yellow page users, users also like more advertisements. On the user demand side, the author uses a multinomial logit model to characterize user demand, and estimates the model following Berry, Levinhson, and Pakes (1995). The number of (classified) advertisements enters into consumer utility as a characteristics of yellow pages. On the advertiser side, he derives a logarithm inverse demand function from a stylized advertiser payoff function characterized by constant returns to consumer impression and advertising congestion. Corresponding to the single-purchasing assumption in traditional discrete choice models, the author assumes that users single-home, and that advertisers multi-home. This

structure of single-homing consumers and multi-homing advertisers echoes the competitive bottleneck model that has been studied extensively in the early theoretical literature. Recall that, with the single-homing user assumption, each platform has a monopoly position to advertisers. An advertiser's decision to join which platforms then becomes a series of separate, independent decisions. In equilibrium, the platform firms - or yellow pages - maximize profit by choosing the number of ads to admit, internalizing the effects of each side's participation on the other side. Rysman (2004) has greatly influenced many subsequent empirical economic studies on media and other two-sided markets. Most of the later works that have a structural econometric component more or less follow it in some way, such as the characteristic approach to modeling media demand, estimation of ad demand using an aggregate demand equation, and its underlying competitive bottleneck structure. In fact, in Rysman (2004), the author meticulously outlays the necessary assumptions for the competitive bottleneck structure, and supports it with stylized facts on the U.S. yellow page markets. Later works sometimes lack the scrutiny of assumptions found in Rysman (2004). In this thesis, I follow Rysman (2004)'s characteristic approach to modeling consumer demand and his approach to estimating advertiser demand with aggregate data. However, I provide very different micro-foundations for the demand functions that I estimate.

The yellow page industry bears some similarities with many broadcasting media and print media industries that have the two-sidedness and are (at least partially) financed through advertising. Nonetheless, unlike in the markets for yellow pages, media consumers often derive utility from content other than advertising, and in fact, they may find advertising a nuisance. Wilbur (2008)'s study on the U.S. television markets is a prominent example that combines a structural econometric model of two-sided markets and high-frequency data on TV audience shares and program characteristics. The underlying market structure is also assumed to be a competi-

tive bottleneck: in any half-hour period, viewers watch only one channel, through which advertisers can deliver their messages. A major contribution of Wilbur (2008) is that, after credibly addressing the endogeneity problem associated with advertising levels, the author finds that viewers dislike ads relative to other non-ad program time. This supports the consumer ad-aversion assumption often made in theory works on two-sided markets, such as in Anderson and Coate (2006), and has important implications for the business model of media platforms as well as social welfare. In short, in a two-sided media market with ad-averse consumers, platforms have to balance between a less-audience situation and a less-advertising one. Using highly detailed data on audience shares, Wilbur (2008) also sheds lights on another internal conflict of two-sided markets: viewers and advertisers are found to have different preferences for program types. He interprets it as evidence that advertisers value attention - or metaphorically “eyeballs” - of viewers differently. It is a leading work in a sizable literature that uses a two-sided structural model to investigate empirical questions in ad-financed media markets. Some other outstanding structural empirical papers on two-sided markets include Jeziorski (2014) on radio and Zhou (2016) on video game consoles.

It is worth noting that when advertising seems to not matter to media consumers, the full two-sided structure of media markets is lost since there is no participation externality for platforms to worry about on the consumer side. A few important studies on newspaper markets in the U.S. and Italy, including Gentzkow (2007), Argentesi and Filistrucchi (2007), Fan (2013), and Gentzkow, Shapiro, and Sinkinson (2014), have adopted a similar consumer ad-neutrality assumption supported with empirical evidence. Nonetheless, a number of leading works in the newspaper studies question the then-prevalent assumption of single-homing consumer in media markets: while it is reasonable to assume consumers only watching one channel or listening to one radio station at a relatively short period of time, it is less reasonable to make the

same assumption about readers of print media in longer time intervals like a year. Gentzkow (2007) develops a discrete choice model that allows newspaper readers to choose two products and is estimatable with individual-level purchase data. Because it does not estimate a micro-founded model of advertising demand, I defer my discussion of this paper to later.

Fan (2013) makes an important contribution by developing a multiple discrete choice model of reader demand and estimating the model using only market-level data. Following a similar intuition in Hendel (1999), the author derives aggregate market penetration of local newspapers as summation of shares of first-choice and second-choice subscribers. She shows that such a market penetration function based on multi-purchasing behaviors satisfies a set of “regularity” conditions so that the powerful methodology of Berry, Levinhson, and Pakes (1995) is applicable. Besides that, another important contribution of Fan (2013) is to endogenize product characteristics in addition to prices and to illustrate the importance of endogenizing product characteristics in merger analyzes. To do that, the author suggests and uses an innovative instrument - overlapping regional demographics of readers - for the endogenous quality problem. In this thesis, I extend Fan (2013)’s model to allow consumers to purchase more than two products at a time. I aid the identification with more detailed data on consumer multi-homing.

The multi-homing reader assumption is critical to modeling advertiser behavior in two-sided markets. Specifically, when both media consumers and advertisers are allowed to multi-home, an advertiser’s decision to join some platforms would become a series of interdependent decisions: since hitting a consumer multiple times may be marginally less effective, an ad placement in one media outlet can influence advertiser payoff in another platform. As a result, a two-sided multi-homing model is usually inconsistent with aggregate (inverse) advertising demand equations - commonly used

in the literature.³ In their study on ideological competition of U.S. newspapers, Gentzkow, Shapiro, and Sinkinson (2014) presents the first structural media market model consistent with two-sided multi-homing. On the reader demand side, they closely follow Gentzkow (2007). On the advertiser side, they derive an equilibrium advertising equation, modeling after Anderson, Foros, and Kind (2016). Their advertising equation embodies the important feature that, to advertisers, the exclusive readership has higher weights than the overlapping one does since it is more profitable for advertisers to appeal to the fresh eyeballs whose attention is not shared otherwise. Using historical data on U.S. newspapers in the early 20 century, they combine estimation and calibration to obtain their empirical results. In particular, they find evidence that exclusive readership is more preferable than overlappers. Along with them, my thesis provides the first direct evidence to support the multi-homing hypotheses in the theory literature. The advertising demand equation in my thesis is consistent with two-sided multi-homing and is identified with using detailed data on magazine ad levels and prices.

Among all media, the magazine markets deserve special attention because the markets often have a full two-sided structure: consumer magazines are usually supported by both circulation and ad revenues (unlike radio and some TV), and both consumers and advertisers do care about participation of the other side (unlike newspapers with empirically-shown ad-neutral readers). Most of the existing recent studies have focused on the German magazine industry due to data availability.⁴ One of the earliest micro-founded empirical models of two-sided magazine markets is developed by Kaiser and Wright (2006), which uses a panel on German magazines for estimation. Kaiser and Wright (2006) derives demand equations of the two sides from a Hotelling

³This statement holds unless one is willing to make some strong assumptions regarding consumer attention, such as constant return to consumer impression, implicitly in Fan (2013).

⁴Some earlier studies on the magazine markets have used U.S. magazine data to perform more “reduced-form”-type analysis. See, for example, Koschat and Putsis (2000a; 2000b), Depken (2004), and Depken and Wilson (2004).

model of consumer and advertiser demands, assuming both sides single-homing.⁵ To perform the analysis, the authors estimate their model using only the top two magazine brands in each major genre. Even after addressing the endogeneity problem using fixed effects and instruments, they find that consumers like the number of ads in magazines. Chandra and Collard-Wexler (2009) builds a similar theoretical model to investigate empirically the effects of mergers in the Canadian newspaper industry. A follow-up paper to Kaiser and Wright (2006) is Kaiser and Song (2009), which only estimates the German magazine reader side a la BLP. Exploiting the panel structure of the data and using standard instruments, Kaiser and Song (2009) find heterogeneous consumer ad preferences towards magazine advertising. For instance, in many genres (e.g., Women’s, Car and TV magazines), consumers seem to appreciate higher advertising ratios. Along with the earlier work, they provide some solid evidence on consumer ad-loving, adding fuels to the heated debate on consumer ad preference in media markets. The empirical evidence presented in this thesis is in contrast to those findings of ad-loving consumers in German magazine markets and ad-neutrality in U.S. newspaper markets. My three-way panel (product-time-year) data gives me more credible exogenous variation: market penetration of magazines is recorded at the metropolitan level while product characteristics do not vary across regions in any given year.

2.3 Related Literature on Modeling Multiple Consumer Choices

This thesis builds upon other empirical models of consumer multi-purchasing behavior. Recall that typical demand estimation with individual-level data or with

⁵The authors also extend their stylized model to account for multi-homing, and the results do not change qualitatively.

market-level data (e.g., Berry, Levinsohn and Pakes (1995)) requires that consumers choose only one product from a set of differentiated products. This empirical assumption has important implications for advertiser behavior and platform competition in two-sided markets - as discussed earlier - and greatly limits the scope of market structures that can be studied. Fortunately, there are at least three groups of flexible choice model that one can use for modeling multiple purchasing behavior. I follow one of the groups - including Hendel (1999) and Fan (2013) - to formulate probabilities (or market shares/penetration) as aggregation of a bunch of mutually exclusive events. In Hendel (1999), the demand for multiple units of an input is modeled as the sum of some task-specific sub-demands that each requires at most one unit of input. In Fan (2013) and this thesis, the probability that any consumer buys a product is the sum of the probabilities that she buys it as her first-choice, second-choice, and etc. The advantage of this approach is its moderate data requirement: such a model is often estimatable and identified with market-level data and some additional micro-moments or market restrictions⁶ and is estimatable using the common BLP-type estimation routine.

In one other line of research, including Gentzkow (2007), Gentzkow, Shapiro, and Sinkinson (2014), and earlier papers in the marketing field, multiple-purchasing behavior is modeled as choosing a bundle of products. In particular, the set of possible choices is defined over all bundles of product, and each bundle has an i.i.d. taste shock. Other parts are identical to a typical multinomial logit or a mixed logit model. One problem associated with this approach is that the cardinality of the set of possible bundles would explode quickly as the number of products increases. Gentzkow (2007), as well as earlier marketing papers, estimates the model using data on individual purchases. Gentzkow (2007) also provides a deep discussion on identification of the

⁶To identify the utility loss from reader multiple purchasing, Fan (2013) assumes that in some markets, local newspapers have monopoly positions and do not face competition from national newspapers. In my thesis, I use more detailed survey data on consumer multi-homing. See Section 6.2.2 for more details.

model with individual-level data. However, Gentzkow, Shapiro, and Sinkinson (2014) applies the model to market-level data, thus compromising identification of some of the parameters.

Finally, a third line of research models consumption of each product as driven by a separate binary choice equation. For example, Augereau, Greenstein, and Rysman (2008) (on R&D decisions) and Hiedemann, Sovinsky, and Stern (2013) (on family long-term care decisions) use multivariate probit models. In those models, consumption of a product is independent of other products, so it implies that each choice has a monopoly position. Specifically, in this kind of model, all cross-price elasticities of products are automatically zero. If one is interested in studying substitution patterns between products and strategic firm behavior, this type of model is not useful. In the thesis, I point out an important connection between the model I use here and the binary model: when consumers are free to choose whatever number of products and there is no diminishing utility from multiple purchases, the multiple discrete choice model becomes a binary choice model. Intuitively, this is when consumers' purchase decisions become independent of the existence of other products, so she only compares the value of a product to the no-purchase option. In Appendix A.1, I present a short formal proof of the above statement.

Chapter 3

Industry Background

The magazine industry is characterized by a large amount of small magazine titles and a small number of giant ones. In 2012, there were 7,390 print magazine brands in the United States, and despite the impact of the Internet, this number has hardly changed since 2004 (National Directory of Magazines 2015).¹ Annually, there are approximately 300 million copies of magazine sold for single issues of magazine. Top 200 magazines account for 85% of all sales while top 50 magazines account for about 54%. By industry standards, magazines brands are usually categorized into genres, or categories, which are the bases for market reporting and analysis. In this paper, I define market segments as genres of magazines, following definitions provided by MPA and the Alliance of Audited Media (AAM).

Magazines subscription constitutes more than 93% of all sales.² This implies that, on average, U.S. households subscribe to between 2 and 3 magazines annually, which is consistent with Hong (2007)'s study based on the Consumer Expenditure Survey that U.S. households spend \$6-\$7 quarterly on magazines. It further indicates that, among magazine subscribers, multiple purchasing, or multi-homing *across genres* is common. Indeed, according to *Survey of the American Consumer* conducted by the

¹Data Source: The Association of Magazine Media (MPA), Kantar Media, GfK MRI and my data set.

²I refer to magazine consumers as readers or subscribers throughout this paper,

media firm GfK MRI, more than 65% of a total of 23,000 surveyed households indicate a second magazine choice. These facts call for a need to incorporate multi-homing behavior into models of consumer choice in magazine markets.

Due to the rise of the Internet, consumers spend less on magazines. Hong (2007) documents that quarterly spending on magazines has dropped 29% on average from 1996 to 2002. Among Internet users, the drop is almost 50%. However, various anecdotes and researches suggest that impact of the Internet on magazines is moderate in comparison to the effects on traditional newspapers.³ This may explain partially why the magazine industry surpassed newspapers to become the second largest advertising market in 2000's (Kantar Media 2012).

However, the magazine industry has remained as a prominent advertising market. First, the magazine industry is mainly supported by advertising revenues to provide valuable media content - whether informational or entertaining - to readers. In 2012, eighty percent of the magazine industry revenue, or \$20 billion, came from selling magazine space to advertisers. Second, the magazines are provide indispensable platforms for advertisers to deliver their messages to the consumers. In fact, up to the year of 2012, the magazine industry had been the second largest advertising market where advertisers have poured their money in order to reach to the magazine consumers. As shown in Figure 3.1., for example, in 2012, 16.2% of all advertising dollars in the United States is spend in consumer magazines while 53.1% is in television, 12% in newspaper, 8.7% in the internet, and 6.2% in radio.

The advertising markets of U.S. magazines are highly special in at least two aspects. First, ratios of advertising pages to non-ad content pages have constantly been around 1:1 over four decades since early 1970's. It is largely the case at the industry level as shown in Figure 3.2 and at the individual product level; for example, see Table 5.1 in Chapter 5. In other words, about fifty percent of magazine pages are

³For example, see Chandra and Kaiser (2016).

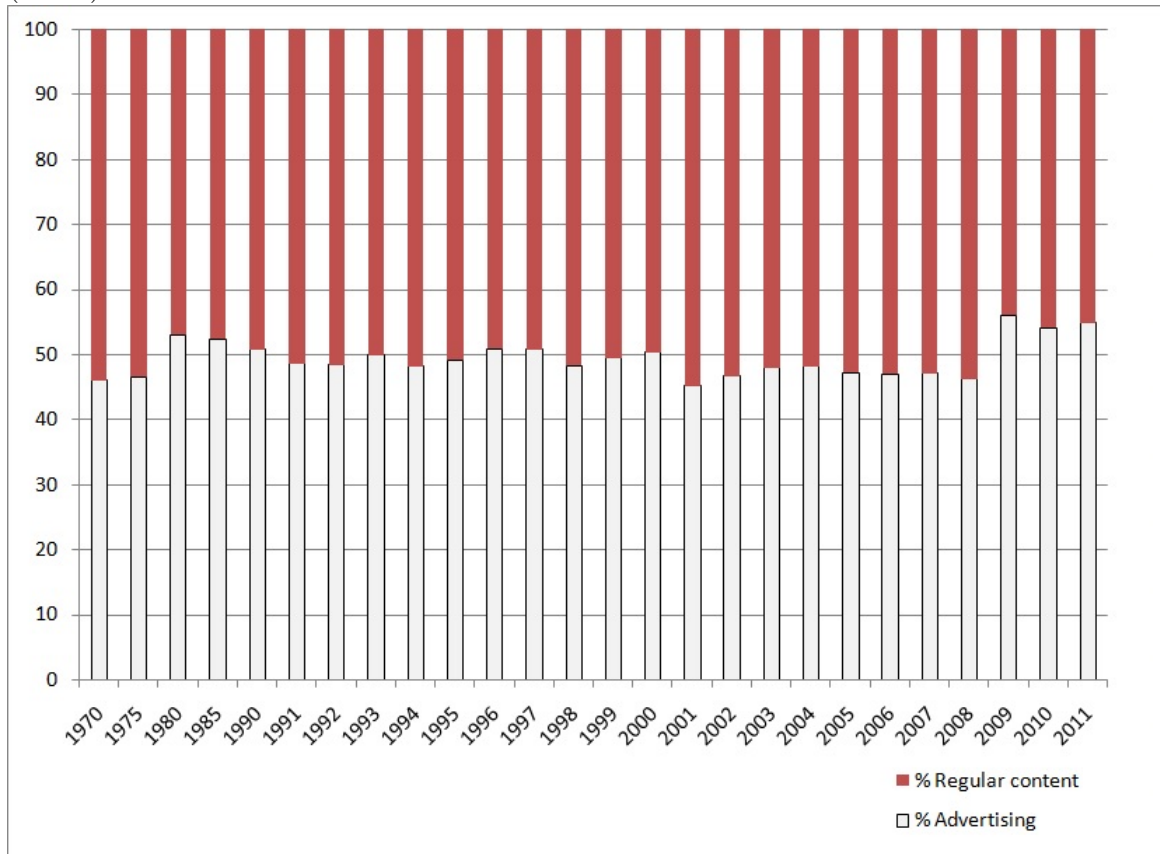
Figure 3.1: Advertising Spending in Media

% of ad dollar spent in media	2008	2009	2010	2011	2012
Magazine	18.1	17.1	16.8	16.1	16.2
Internet	6.7	7.9	7.6	9.1	8.7
TV	47.2	49.3	51.1	51.2	53.1
Radio	6.8	6	6.1	6	6.2
Newspaper	16.1	14.8	13.6	12.9	12

(Source: Kantar Media)

advertisements. This number is high in comparison to other media industries like television and radio. Wilbur (2008) reports an average of a little less than five minutes of commercials in a thirty-minute interval, or about 17 percent of total program time. On the internet, it is also hard to imagine that advertising spans half of a webpage. Second, large magazine platforms - those that have many subscribers - usually set very high prices of advertising. For the top 20 U.S. magazines, advertisers pay approximately \$280,000 per page of advertising on average. Advertising prices are still considered very high at the per-reader level. For instance, as reported in Chapter 5, the mean ad price per subscriber in my sample of 34 major magazines is \$0.15, or \$150 for CPM (i.e., cost per thousand consumers). At the meantime, the CPM for Super Bowl ads is around \$30 - an industry-wide high in U.S. television.

Figure 3.2: Ratio of Ad Pages vs. Content Pages in the U.S. Magazine Markets (MPA)



Chapter 4

The Model

4.1 Overview of the Model

Magazines markets exhibit characteristics of a two-sided market: magazine firms or platforms sell media content to ad-sensitive readers and in turn charge advertisers for using their space for advertising. In this chapter, I describe a structural model of two-sided magazine markets, which includes characterization of the (reader) demand for magazines, the demand for magazine advertising, and pricing decisions of magazines. The model is set up for estimation with newly available data on metropolitan-level magazine circulation, consumer ranking of magazines, and advertising prices and quantities. It incorporates features of the data and of the magazine industry that are most relevant and important to this study - including consumer preference heterogeneity, multiple purchasing or multi-homing, and platform ad pricing that is based on the composition of consumers. The demand functions in Section 3.2 and Section 3.3 of this chapter are used directly in the estimation, and they also give testable implications for the data. Magazine pricing equations in Section 3.4 complete the characterization of market equilibrium. Combining with demand-side estimates, I also use the pricing equations to infer marginal costs of magazine production and

advertising. However, the structural model in this chapter is kept intentionally as general as possible, so minimal functional form assumptions are made for the analysis. Full specifications of this model is presented in Chapter 6 along with details of estimation and identification of the model.

4.2 Demand for Magazines

I model consumer demand for magazines as a multiple discrete choice problem among differentiated products. The metropolitan-level demand for magazines is derived from aggregation of multiple discrete choice probabilities of consumers. The model is needed to explain consumer multi-homing in the data and to predict the level of exclusive readerships on each platform. Based on Hendel (1999) and Fan (2013), my model allows consumers to purchase multiple products - or magazines - at a time, and the incremental utility of a product diminishes after each choice of product. Using a unique data-set on consumer ranking of purchased magazines, I add richness to identification of this type of model. To model consumer utility from choosing a single product, I take the standard characteristic approach as in a common multinomial discrete choice model. I then describe, first, how consumers can choose up to four magazines; second, the expression of an individual's choice probability; and lastly, aggregation of individual choice probabilities into market penetration of products. I defer details of full econometric specifications of the reader demand model to Section 6.1.1 of Chapter 6.

To begin with, consider a consumer i in a metropolitan area c , deciding whether or not to subscribe to some national-brand consumer magazines, indexed $j \in \{1, \dots, J\}$, in year t .¹ Because all the magazines are national-brand, consumers in all metropolitan areas have the same choice set. A consumer can also choose the “outside” option,

¹I assume that all purchases are made via subscription. As discussed in Section 2, subscription accounts for 93% of all magazine circulation.

indexed $j = 0$, which is to not purchase any magazine in the choice set. Consumer i 's conditional indirect utility from purchasing a single magazine j is:

$$u_{ijct} = \delta_{jct}(\mathbf{p}^s, \mathbf{a}) + \vartheta_{ijct}(\mathbf{p}^s, \mathbf{a}) + \varepsilon_{ijct}, \quad (4.1)$$

where δ_{jct} is the mean utility for magazine j in market ct ; and the ϑ_{ijct} and ε_{ijct} terms are the individual-specific taste components for j , capturing any deviation from the regional mean. \mathbf{p}^s and \mathbf{a} are vectors of magazine subscription price and ad level. As described in detail in Section 6.1.1, ϑ_{ijct} contains interactions between product characteristics and individual tastes, while ε_{ijct} is the pure idiosyncratic error term. As a consumer observes her utility - including all three terms for each choice - upon making a decision, firms and econometricians do not observe ε_{ijct} and the random components of ϑ_{ijct} . As in a standard multinomial logit model, I assume ε_{ijct} to be i.i.d. with Type I extreme value distribution before any further specifications of the utility function.

I now turn to describe how consumers make multiple purchasing decisions. The following is a heuristic description of the decision-making process of consumers. Specifically, each household may choose to subscribe to multiple magazines at a time, and for each magazine title, they would purchase no more than one unit of subscription. One's first(-best) choice is the magazine title that yields the highest utility among all available product choices. For the second(-best) purchase, the incremental utility from subscribing to any magazine is given by $u_{ijct} - \kappa$; that is, the conditional indirect utility of product j decreased by a constant, κ . One's second choice is then the product that yields the highest utility among the *unchosen* products. Similarly, for the n^{th} choice, the incremental utility is $u_{ijct} - \kappa(n - 1)$, or utility from consuming product j decreased by $\kappa(n - 1)$. Consumers would keep subscribing to magazines until the no-purchase option yields the highest utility or until they reach four sub-

scriptions. This assumption is mainly due to data limitations and the fact that 98% of magazine subscribers buy no more than four magazines *across genres*.

The diminishing utility parameter κ captures any utility loss from subscribing to more than one magazine of the same category at the same time. Inclusion of this parameter is motivated by the idea that consuming, say, two same-type magazines at the same time yields less utility than the sum of two stand-alone utilities. This can be attributed to scarcity of readers' leisure time. Alternatively, the utility loss can arise from overlapping content of any two same-genre magazines. For example, two business magazines may cover similar news event and data but differ in their commentating and other editorial content. So, consumers would benefit from reading the news event only *once* while valuing any distinct interpretation of the same news in two magazines. Favoring the latter story, Gentzkow (2007) and Gentzkow, Shapiro, and Sinkinson (2014) include a utility loss parameter that is product-pair-specific.² Due to data limitation, I assume κ to be constant, favoring the first type of interpretation. The same parameter is also in Fan (2013). However, while I expect κ to be positive, I do not put any restriction on it throughout the econometric implementation.

4.2.1 Choice Probability and Market Penetration Function

Given the above description of the consumer choice problem, we can see that the events that a magazine j are i 's first(-best), second(-best), third(-best), and fourth(-best) choices are mutually exclusive since any consumer would purchase no more than one subscription per magazine title. Because these choice events are mutually exclusive, in general, the probability that consumer i *ever* subscribes to magazine j

²It is worth nothing that, in Gentzkow (2007) and Gentzkow, Shapiro, and Sinkinson (2014), there can also be a utility gain for any bundle of products - they do not put any restriction on the sign of this utility parameter - similar to κ in my model - *a priori*.

(with ct suppressed) is given by:

$$\Pr\left(u_{ij} \geq \max_{h \in \mathcal{J}} u_{ih}\right) + \sum_{n=2}^4 \sum_{j^{(1)} \dots j^{(n)} \neq j} \Pr\left(u_{ij^{(1)}}, \dots, u_{ij^{(n)}} > u_{ij} \geq \max_{h \neq j^{(1)} \dots j^{(n)}} u_{ih}, u_{ij} - \kappa(n-1) \geq u_{i0}\right); \quad (4.2)$$

where the first term is the probability j is i 's first(-best) choice, and the subsequent terms are the probabilities of j being the n^{th} (-best) choice. For illustration, consider the case of $n = 2$ (i.e., consumers can buy at most 2 products) assumed in Fan (2013). There, consumer i 's choice probability of buying magazine j is:

$$\Pr\left(u_{ij} \geq \max_{h \in \mathcal{J}} u_{ih}\right) + \sum_{j^{(1)} \neq j} \Pr\left(u_{ij^{(1)}} > u_{ij} \geq \max_{h \neq j^{(1)}} u_{ih}, u_{ij} - \kappa \geq u_{i0}\right). \quad (4.3)$$

Let Φ_{ij}^1 denote the probability j is i 's first(-best) choice, $\Phi_{ij}^{1,1}$ the probability j is i 's first choice with utility decreased by κ once, and $\Phi_{ij-j^{(1)}}^{1,1}$ the probability j is i 's first choice given that $j^{(1)}$ is already chosen. Notice that $\Phi_{ij-j^{(1)}}^{1,1}$ is also the probability of j being *at least* the second choice given that $j^{(1)}$ may be the first. Conceptually, the probability of j being *at least* the second choice given that $j^{(1)}$ may be the first consists of the probabilities of two mutually exclusive events: the event that j is *exactly* the second choice given that $j^{(1)}$ is the first and the event that j is i 's first choice (even in comparison to $j^{(1)}$). Therefore, using our notation, $\Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1}$ is the probability of j being *exactly* the second choice given that $j^{(1)}$ is the first. Notice that, up to this point, I only use mutual exclusivity of these probability events. Generally, for a discrete choice model, the probability consumer i purchases magazine j with the restriction $n = 2$ can be written as:

$$\Phi_{ij}^1 + \underbrace{\sum_{j^{(1)} \neq j} \left(\Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1}\right)}_{\Phi_{ij}^2}. \quad (4.4)$$

Similarly, I present the (recursive) expression for the case of $n = 4$ in the Appendix A.2. Recall that the assumption $n = 4$ is mainly due to the empirical observation that 98% of magazine subscribers buy no more than 4 magazines *across genres*. While it is not difficult to derive the choice probability expression for the case of $n = 5$, to establish some regularity conditions for that case requires additional work.

Proceeding to write down functions for the probabilities, I now use the assumption that ε_{ijct} is i.i.d. with Type I extreme value distribution. Under the distributional assumption, I can write Φ_{ij}^1 , $\Phi_{ij-j^{(1)}}^{1,1}$, and $\Phi_{ij}^{1,1}$ in familiar logit terms:

$$\begin{aligned}\Phi_{ij}^1 &= \frac{e^{\delta_j + \vartheta_{ij}}}{1 + \sum_j e^{\delta_h + \vartheta_{ih}}}, \\ \Phi_{ij-j^{(1)}}^{1,1} &= \frac{e^{\delta_j + \vartheta_{ij}}}{e^\kappa + \sum_{h \neq j^{(1)}} e^{\delta_h + \vartheta_{ih}}}, \\ \Phi_{ij}^{1,1} &= \frac{e^{\delta_j + \vartheta_{ij}}}{e^\kappa + \sum_j e^{\delta_h + \vartheta_{ih}}},\end{aligned}$$

which are used in the estimation. Notice that, when the incremental utility from consuming any product is decreased by κ , we have the value of the outside option increased from 0 to κ : the outside option becomes more attractive. More details are organized in Appendix A.2. In short, we use mutual exclusivity of the probability event and the logit-error assumption to write down expressions of an individual's choice probability.

Given the above characterization of individual choice probabilities, the market demand for product j is then aggregation of each individual's choice probability of subscribing to magazine j . Specifically, the market penetration function of j is:

$$s_j(\delta, \kappa) = \int \int \left(\Phi_{ij}^1(\delta, z, \nu) + \sum_{n=2}^4 \Phi_{ij}^n(\delta, \kappa, z, \nu) \right) dP_\nu dP_z; \quad (4.5)$$

In other words, integration of individual choice probabilities of a product j over the empirical distribution of consumer demographics and the distribution of unobserved consumer taste heterogeneity yields the market penetration of j . In practice, simulation is used to calculate the integral, which does not have a close-form presentation analytically. Notice also that the term "market share" has been avoided since market shares sum up to 1, which is no longer necessary in this case exactly due to consumer multi-homing.

As discussed in detail in Chapter 5, I observe in the data not only sales or market penetration of each magazine at the aggregate level but also the proportion of subscribers who purchase j as their first to fourth choices. So I can match each of $\Phi_j^m = \int \int \Phi_{ij}^m dP_v dP_z$ to corresponding moments in the data. It is useful to notice that $\Phi_{ij}^1(\delta, z, \nu)$, the probability of j being the first choice, does not depend on the utility loss parameter κ , while $\Phi_{ij}^n(\delta, \kappa, z, \nu)$ does for $n = 2, 3, 4$. This difference provides the key source for identification of κ .

4.2.2 Further Discussion on κ and the $n = 4$ Assumption

For the oligopoly analysis, it is actually essential to assume that $\kappa > 0$ and $n < J$. I show that without these assumptions, the multiple discrete choice model described in earlier subsections collapses to a binary choice problem, unsuited for any oligopoly analysis. The proof is quite general, for it only requires mutual exclusivity of the probability events and no distributional assumption. Intuitive, when $\kappa = 0$ and $n = J$ - that is, consumers is free to choose any number of products without any utility loss from multiple purchases - consumers just say yes or no by comparing each product to the no-purchase option only. In particular, one's decision whether or not to purchase a product is independent of the value of any other product. When the choice model becomes binary, it implies that each product has a monopoly position: prices are strategically independent, and all cross-price elasticities are zero. In short,

when $\kappa = 0$ and $n = J$, the demand model becomes inappropriate for any study of firms' strategic behavior. I establish formally the interesting result in Appendix A.1.

4.2.3 Regularity Conditions for the Reader Demand

As in the original Berry, Levinhson, and Pakes (1995), it is important to establish the invertibility of $s_j(\delta)$, the market penetration function in this case of a multiple discrete choice model. The invertibility of the market penetration function means an one-to-one relationship between any product j 's market penetration, $s_j(\delta)$, and its mean product utility, δ_j . In other words, in this model when media consumers may choose more than one product to consume, higher market penetration of a product should still imply a higher mean consumer utility level, and vice versa - as in a standard multinomial logit model with single-purchasing consumers.

Following Berry, Levinhson, and Pakes (1995) and Fan (2013), I verify the invertibility of $s_j(\delta)$ in Appendix A.2. In short, it boils down to check the following four regularity conditions for $s_j(\delta)$ in this multiple discrete choice model:

$$(C.1) \quad \partial s_j / \partial \delta_j < s_j,$$

$$(C.2) \quad \partial s_j / \partial \delta_j > 0,$$

$$(C.3) \quad \partial s_j / \partial \delta_h < 0 \text{ for } h \neq j,$$

$$(C.4) \quad \sum_{h=1 \dots J} \partial s_j / \partial \delta_h > 0.$$

In addition to being central to our econometric implementation, these regularity conditions carry very important economic meanings. Conditions (C.1) and (C.2) mean that the market penetration function of any product should be increasing in its own quality; however, the marginal effect of quality on market penetration (or sales) cannot be "too large." Specifically, no marginal change in quality can double its penetration. Condition (C.3) states that the market penetration of a product

should be decreasing in any other product's quality. This insures that products are substitutes. Condition (C.4) means that j 's own quality effect dominates all other cross-product quality effects.

While the regularity conditions seem natural economically, it is non-trivial to verify these regularity conditions in the context of consumer multiple purchasing. For instance, consider condition (C.3) which states that any cross-product quality effect should be negative. In the context of traditional discrete choice models, it is equivalent to state that an increase in any other product's quality would decrease the probability that any consumer buys j (as her first best choice). In the standard logit model, it is straightforward to show that $\partial s_j / \partial \delta_h = -s_j s_h < 0$. However, in our case of the multiple discrete choice model, we have:

$$\frac{\partial s_j}{\partial \delta_h} = \int \int \left(\frac{\partial \Phi_{ij}^1}{\partial \delta_h} + \frac{\partial \Phi_{ij}^2}{\partial \delta_h} + \frac{\partial \Phi_{ij}^3}{\partial \delta_h} + \frac{\partial \Phi_{ij}^4}{\partial \delta_h} \right) dP_v dP_z.$$

While it is obvious that $\partial \Phi_{ij}^1 / \partial \delta_h < 0$ since Φ_{ij}^1 is in a familiar (mixed) logit form, it is not immediately clear that $\partial \Phi_{ij}^2 / \partial \delta_h < 0$. Indeed, as shown in Appendix A.2,

$$\frac{\partial \Phi_{ij}^2}{\partial \delta_h} = \Phi_{ij}^{1,1} \Phi_{ih}^{1,1} + \sum_{j^{(1)} \neq j, h} \left(-\Phi_{ij-j^{(1)}}^{1,1} \Phi_{ih-j^{(1)}}^{1,1} + \Phi_{ij}^{1,1} \Phi_{ih}^{1,1} \right);$$

where the first term is positive, and the second (summation) term is negative. Without further quantification, the sign of $\partial \Phi_{ij}^2 / \partial \delta_h$ is ambiguous *a priori*. In other words, in this case when market penetration of a product consists of shares of first-choice to fourth-choice consumers, one cannot easily conclude that the cross-product effect of quality is negative. This is because higher quality of other products do not necessarily reduce the share of second-choice or third-choice consumers. Similar challenges occur in showing other conditions as well. To notice, the proof that I present in Appendix A.2 often utilizes the functional form assumption of logit errors. In addition, my proof for the case of $n = 4$ does not immediately extend to any case of $n > 4$; for

instance, to verify those conditions hold for $n = 5$ requires substantial additional work. However, I believe the key ingredients in my proof - as well as in Fan (2013) - should be useful to generalize our results.

4.2.4 Exclusive Readership

To examine equations (4.4) and (4.5) more closely, we see that this model implies that magazines with larger penetration s_j (and thus larger sales $\mathcal{M}s_j$) should have more first-choice subscribers; that is, those who rank magazines as their first choice. To see that, first notice Φ_{ij}^1 is increasing in δ_j , the mean utility for j . By the invertibility of the market penetration function shown in Appendix A.2, a large penetration s_j implies a large δ_j . It follows immediately that a large penetration implies large Φ_{ij}^1 , hence a large Φ_j^1 . Intuitively, magazines that sell a lot of subscriptions have higher mean consumer valuations, which lead to many of the subscribers regarding them as their first choice. This is the first important link that I examine with the data. Fail to see such a relationship in the data would lead to immediate rejection of the model.³ In Section 5, I present data patterns in support of this relationship.⁴

Next, it is useful to write down the following expression for the proportion of exclusive readers on platform j - consumers who subscribe to magazine j only:

³Of course, this relationship is only a necessary condition.

⁴It is natural to wonder whether larger market penetration implies, say, a higher proportion of second-choice consumers. While such a relation seems very intuitive, I am unable to verify that mathematically. For instance, it requires that $\frac{\partial \Phi_{ij}^2}{\partial \delta_j} > 0$. However, we can see that:

$$\begin{aligned} \frac{\partial \Phi_{ij}^2}{\partial \delta_j} &= \sum_{j^{(1)} \neq j} \left(\Phi_{ij-j^{(1)}}^{1,1} (1 - \Phi_{ij-j^{(1)}}^{1,1}) - \Phi_{ij}^{1,1} (1 - \Phi_{ij}^{1,1}) \right) \\ &= \sum_{j^{(1)} \neq j} \left(\Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1} \right) \left(1 - \Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1} \right); \end{aligned}$$

in which, the sign of the second term - $\left(1 - \Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1} \right)$ - is ambiguous. Similarly, the signs of $\frac{\partial \Phi_{ij}^3}{\partial \delta_j}$ and $\frac{\partial \Phi_{ij}^4}{\partial \delta_j}$ can be ambiguous as well.

$$\tilde{\tau}_j(\delta, \kappa) = \int \int (\Phi_{ij}^1(\delta, z, \nu) \cdot \Phi_{i0-j}^{1,1}) dP_\nu dP_z, \quad (4.6)$$

where $\Phi_{i0-j}^{1,1} = \frac{e^\kappa}{e^\kappa + \sum_{h \neq j} e^{\delta_h + \vartheta_{ih}}$. In words, magazine j 's exclusive eyeballs are consumers who choose j as the first choice and choose no other magazines. The proportion/number of exclusive eyeballs on platform j is thus a function of the percentage/number of buyers who regard the product as the first choice. This relationship provides the second important link that I shall revisit when I formulate the demand for advertising in later sections. In addition, for each magazine j , equation (4.6) is used to predict missing data on exclusive readerships, which are a key variable to explain ad pricing of magazines.

4.2.5 Discussion on the Modeling Choice

The reader demand model described in this chapter closely follows Fan (2013), in which a newspaper reader demand model with the restriction of $n = 2$ is used for estimation with data on newspaper sales in the U.S. but no consumer survey data on the extent of multi-homing behavior. As discussed in detail in Chapter 2, this approach differs from another major method to model consumer multiple purchasing, employed by Gentzkow (2007) and Gentzkow, Shapiro, and Sinkinson (2014). In Gentzkow (2007), consumers can choose among bundles of product. The conditional indirect utility from consuming each bundle has an i.i.d. error, which can be hard to justify structurally. In addition, the cardinality of a choice set is given by $2^J + 1$ with J being the number of available products, which grows much faster than the speed that J increases. Moreover, in order to match moments, one needs to have data on market shares of each bundle - instead of shares of each product - when individual-level data are not available. Gentzkow (2007) uses individual-level data to study the choice problem of two major U.S. newspapers. In Gentzkow, Shapiro, and Sinkinson

(2014), full identification is not achieved partly due to issues of their market-level data. Therefore, one advantage of the reader demand model that I use in this thesis is that it can handle the case with many products. In fact, when n is a small number, such as $n = 2$ in Fan (2013), to write down the expression for market penetration of a product is straightforward. In my case, it is slightly more complicated but still quite doable. More importantly, another advantage is the relatively mild data requirement of this model. As I discuss in the identification section, Section 6.2, my data on magazine circulation and consumer survey are sufficient for the purpose of identification. In fact, if one is willing to make some stronger assumption, market-level sales or circulation data alone may well be sufficient. For example, Fan (2013) has newspaper circulation data, and she assumes that in regions where there is only one local newspaper, that local newspaper has a monopoly position and faces no competition from national-brand newspapers.

One advantage of Gentzkow (2007) and Gentzkow, Shapiro, and Sinkinson (2014) is that they can allow the bundle interaction term to be pair- or bundle-specific, and they do not have to impose any assumption regarding substitutability or complementarity between products. With more detailed data available, I can revise my model to have a magazine pair-specific κ . However, the substitutability assumption - that is, $\kappa > 0$ - is essential for my model. As a result, the model here cannot handle satisfactorily a situation in which consumers choose among complementary products.

4.3 Demand for Advertising

I model advertisers through their aggregate demand for advertising on magazine platform j . The model is stylized to best suit the data on aggregate magazine advertising prices and quantities. But nonetheless, it captures consistently important features of the magazine industry. In particular, it echoes the formulation in the theoretical

literature on two-sided market competition with multi-homing. Estimating the advertiser demand, I provide the first direct evidence that media advertising prices reflect advertisers' differential valuation of exclusive and non-exclusive eyeballs. The latter hypothesis is constructed in the theoretical literature to confront real-world puzzles that platforms with large audiences often charge higher per-audience ad prices.

Consider a continuum of advertisers indexed by a . Each advertiser can place one ad on platform j in order to reach the subscribers. Each advertiser is a monopoly producer, so there is no product competition among advertisers. The first impression of a subscriber is worth λ_1 to advertisers, and all subsequent impressions are worth λ_2 , with $\lambda_1 > \lambda_2$. For any given set of demand shifters, advertisers are ranked in terms of their willingness to pay in descending order: the reservation price of advertiser a is $p_j(a)$. The advertiser with the lowest reservation price is willing to pay λ_1 per exclusive subscriber. Therefore, an advertiser is only willing to pay λ_2 for each pair of non-exclusive eyeballs. I assume the following linear inverse demand function for advertising in platform j :⁵

$$p_{jt}^a = \lambda_1 \tau_{jt}^e + \lambda_2 \tau_{jt}^o + \lambda_3 a_{jt} + y_{jt}; \quad (4.7)$$

where p_{jt}^a is magazine j 's per-page advertising price; τ_{jt}^e and τ_{jt}^o represent the level of exclusive readers and non-exclusive readers, respectively; y_{jt} captures the total effect of other demand shifters, which full specification is deferred to Chapter 6. By definition, $\tau_{jt}^e = \mathcal{M} \tilde{\tau}_{jt}$ and $\tau_{jt}^o = N_{jt}^s - \tau_{jt}^e$, with N_{jt}^s being the subscription level of magazine j . Finally, a_{jt} is the ad quantity.

Substituting $\tau_{jt}^o = N_{jt}^s - \tau_{jt}^e$ into equation (9) yields the following equation:

$$p_{jt}^a = \tilde{\lambda}_1 \tau_{jt}^e + \lambda_2 N_{jt}^s + \lambda_3 a_{jt} + y_{jt}; \quad (4.8)$$

⁵I assume that advertisers have always reached non-exclusive subscribers elsewhere besides J platforms. This implies whether advertisers multi-home or not, they never capture the first impressions of non-exclusive consumers.

where $\tilde{\lambda}_1 = \lambda_1 - \lambda_2$ while all other coefficients have the same interpretations as in (9). If consumers' first impressions are more valuable than subsequent ones, then I expect $\tilde{\lambda}_1 > 0$. Failure to include data on τ_{jt}^e when $\tilde{\lambda}_1$ is significantly greater than zero would create an upward bias in the estimate of λ_2 and would result in incorrectly estimating a platform's pricing capability.

This equation captures important insights in the theoretical two-sided market literature: when consumers multi-home in media markets and returns to consumer impression drop after the first impression, platform advertising prices should reflect advertisers' differential valuation of consumers. In particular, ad prices should reflect the demand-side phenomenon that advertisers value exclusive eyeballs of single-homing consumers more than the non-exclusive ones. It becomes useful in explaining the so-called "ITV premium puzzle" or "Fox News puzzle" - media platforms with larger penetration have greater per-audience ad prices - that contribute to motivate the somewhat nascent literature on multi-homing in two-sided markets. I discuss more in Section 4.4, when I lay out the platform maximization problem.

One caveat about this linear advertiser demand function is that it can be hard to interpret under some circumstances. For instance, when a magazine j delivers to zero consumer, or $N_{jt}^s = 0$ and $\tau_{jt}^e = 0$, the advertising price is still positive. It then leaves one wondering why any advertiser wants to pay if her ad would reach to no consumer. However, this linear demand model can be seen as a good approximation of the "true" underlying model away from the extreme point with $N_{jt}^s = 0$ and $\tau_{jt}^e = 0$. Wilbur (2008) uses a similar linear demand function for advertising except that there is no notion of consumer multi-homing and exclusive consumers.

4.4 Magazine Platforms

Magazine platforms are assumed to set their product characteristics before prices. This is a common assumption in the literature. So, I view product characteristics as exogenous to both the subscription price and the number of ads. Magazines choose subscription prices and ad levels simultaneously. Observing subscription prices and anticipating the amount of advertising on each platform, consumers make their subscription decisions as the result of utility maximization. Given both the subscription level and the composition of reader base, each magazine chooses the amount of advertising to include, which yields correspondingly the per-page advertising price. At the same time, advertisers get admitted based on their willingness to pay.

In reality, consumers often do not observe the annual total number of ads when they decide to purchase a magazine subscription. However, they should rationally anticipate the amount of ads in each magazine, and in equilibrium, their expectation should be consistent with the realized ad level. Therefore, when setting the number of ads to carry, platforms need to internalize the effects of advertisers' participation on subscribers, which in turn affect the profitability of the advertising market.

Let $N_{jt}^s \equiv \sum_c \mathcal{M}_{ct} s_{jct}$. Magazine j 's profit maximization problem is therefore:

$$\max_{p_{jt}^s, a_{jt}} \pi_{jt} = \underbrace{N_{jt}^s(\mathbf{p}^s, \mathbf{a})(p_{jt}^s - c_{jt}^s)}_{\text{circulation profit}} + \underbrace{a_{jt}(p_{jt}^a(a_{jt}, \tau_{jt}^e(\mathbf{p}^s, \mathbf{a}), N_{jt}^s(\mathbf{p}^s, \mathbf{a})) - c_{jt}^a)}_{\text{advertising profit}}; \quad (4.9)$$

where c^s and c^a are j 's marginal costs associated with circulation and providing advertising space, respectively. c^s reflects the marginal cost of physical production of each copy, while c^a captures marginal costs of advertising, such as costs of production and sales efforts. In words, magazine j 's total profit comes from both selling and delivering copies to subscribers and selling advertising pages to advertisers.

In this model, advertising prices are charged on a per-page base. This is largely motivated by the practice of charging ads by column-inch in the magazine industry.

I also assume that the marginal costs of advertising are incurred on a per-page base. Similar assumptions can be found in other structural studies of media markets, such as in Rysman (2004), Wilbur (2008), and Fan (2013). At the same time, standard models of two-sided media markets (e.g., Anderson and Coate (2005); Anderson et al. (2014)) assume ad pricing on a per-page per-consumer base. Their motivation is mainly theoretical: as noted by Armstrong (2005), and White and Weyl (2014), per-consumer ad prices avoid complexities in solving theoretical two-side market models.

The Nash equilibrium in magazine markets is defined implicitly by a system of magazines' first-order conditions from the optimization problem:

$$N_{jt}^s + \frac{\partial N_{jt}^s}{\partial p_{jt}^s} (p_{jt}^s - c_{jt}^s) + \left(\frac{\partial p_{jt}^a}{\partial \tau_{jt}^e} \frac{\partial \tau_{jt}^e}{\partial p_{jt}^s} + \frac{\partial p_{jt}^a}{\partial N_{jt}^s} \frac{\partial N_{jt}^s}{\partial p_{jt}^s} \right) a_{jt} = 0, \quad (4.10)$$

and

$$\frac{\partial N_{jt}^s}{\partial a_{jt}} (p_{jt}^s - c_{jt}^s) + a_{jt} \left(\frac{\partial p_{jt}^a}{\partial \tau_{jt}^e} \frac{\partial \tau_{jt}^e}{\partial a_{jt}} + \frac{\partial p_{jt}^a}{\partial N_{jt}^s} \frac{\partial N_{jt}^s}{\partial a_{jt}} \right) + p_{jt}^a + a_{jt} \frac{\partial p_{jt}^a}{\partial a_{jt}} - c_{jt}^a = 0, \quad (4.11)$$

where $\frac{\partial N_{jt}^s}{\partial p_{jt}^s} = \sum_c \mathcal{M}_{ct} \frac{\partial s_{jct}}{\partial p_{jt}^s}$. I then plug (4.10) into (4.11).

Again, the First-Order Conditions (FOCs) from equations (4.10) and (4.11) implicitly define the equilibrium subscription prices and ad levels of platforms. Notice that in (4.10), the third term captures the effect of a subscription price change on platform j 's profitability in the advertising market, due to the two-sidedness of magazine markets. In equilibrium, subscription prices have indirect effects on ad prices through their impacts on both the subscription level and the composition of subscribers. Without the third term, (4.10) is the FOC in standard Bertrand games. Similarly, in Equation (4.11), the first term is the effect of a change in the ad level on the subscription market. The second term captures the feedback effects of a change in the ad level on the advertising market through numbers of subscribers and exclusive subscribers. Without these terms, equation (4.11) is the standard monopolist FOC.

Equation (4.11) implies that there is no direct price competition in the advertising market, although platforms compete indirectly for advertisers via N_{jt}^s and τ_{jt}^e .⁶

The FOCs - in conjunction with the demand-side equations and observations - are useful to help us understand the empirical puzzle that larger platforms charge greater ad prices per-audience. In particular, the demand functions imply that platforms with larger penetration also have more exclusive eyeballs. Since platform ad prices reflect not only the subscription level but also the composition, one would see larger platforms charge higher ad prices even after accounting for their audience base.

Following BLP (1995) and subsequent papers, I derive the following equations to infer marginal costs of each magazine. Let \mathcal{O} be the ownership matrix with elements $\mathcal{O}(h, j) = 1$ if magazine h and j have the same publisher, and zero otherwise. Let ∇_p^s be a matrix containing all of the first partial derivatives of penetration with respect to subscription prices, with elements $\nabla_p^s(h, j) = \frac{\partial N_j^s}{\partial p_h^s}$. Similarly, denote Let ∇_p^r with elements $\nabla_p^r(h, j) = \frac{\partial \tau_j^e}{\partial p_h^s}$. Define ∇_a^s and ∇_a^r similarly for ad pages. So the mark-ups can be computed using the following FOCs written in matrix forms:

$$\mathbf{p}^s - \mathbf{c}^s = -(\lambda_2 \mathbf{a} + (\mathcal{O} * \nabla_p^s)^{-1}(\mathbf{N}^s + \tilde{\lambda}_1(\mathcal{O} * \nabla_p^r))); \quad (4.12)$$

and

$$\mathbf{p}^a - \mathbf{c}^a = (\mathcal{O} * \nabla_a^s) \cdot (-(\mathcal{O} * \nabla_p^s)^{-1}(\mathbf{N}^s + \tilde{\lambda}_1(\mathcal{O} * \nabla_p^r))) + (\mathcal{O} * \nabla_a^r) \cdot \mathbf{a} - \lambda_3 \mathbf{a}. \quad (4.13)$$

where I plug the FOC (4.12) into (4.13).

⁶Although the advertiser demand function is in the same spirit of recent theoretical advancements; i.e., Anderson, Foros, and Kind (2016). It nonetheless implies monopoly advertising markets while Anderson, Foros, and Kind (2016) derives their pricing equation based on a Bertrand-type model. A model that allows both direct price competition for advertisers and consumer multi-homing and that is estimable with aggregate data is yet to be developed.

Chapter 5

Data

5.1 Data Description

I estimate the model using new data on magazine circulation at the MSA level, magazine characteristics, consumer rankings of magazines, and household demographics from four main sources. On the reader demand side, I observe circulation of magazines in nearly all MSAs in the U.S. between 2003-2012. For a subset of magazines, I also observe aggregate percentages of subscribers rank a magazine as their first to fifth choice. I match these penetration and circulation data with product-level magazine prices, ad pages, and other attributes to create a panel of 34 magazines for 10 years. The magazines are the major products in six different genres defined by the Association of Magazine Media (MPA) and the Alliance of Audited Media (AAM). Because circulation is at the regional level while product attributes of a magazine do not vary across regions, I include data on household demographics to explain geographical variation in circulation of a same magazine. To estimate the advertiser side of my model, I use product-level data on advertising pages, prices and other characteristics for the panel of 340 magazine/years. For robustness checks, I use a panel of 640 magazine/years, including magazines with missing sales information. I

include more details on magazine titles and data sources in the Appendices.

I collect detailed magazine circulation data from the 2003-2012 *Magazine Market Coverage* reports administered by the AAM, the organization that audits U.S. print media circulation and other related information.¹ For each magazine, I observe magazine circulation and penetration in each MSA for 10 years. By definition, market penetration is calculated by dividing total circulation in a region by the number of households. The reports also provide each region's number of households, which is the measurement of market size used in print media industries and in the literature.

Magazine subscription prices, ownership and frequencies of publication come from individual audit reports on magazines from the AAM. In particular, the annual subscription price of a magazine is the reported average of subscription prices paid by all subscribers, accounting for discounts and promotions. It is therefore not the listed price. Frequencies of publication include the number of special issues and supplements in addition to regular issues. Content page numbers come from MA-focus Media, a media research firm that systematically collects page information on major U.S. magazines. Advertising pages and rates come from the Publishing Information Bureau (PIB) affiliated with the Kantar Media. For each magazine/year, I observe the total number of advertising pages and revenues. I calculate the average per-page advertising price by dividing total ad revenues with ad page numbers.

Table 3 reports the summary statistics for magazine sales and attributes. The magazines of this study are large platforms in terms of reader base: an average magazine has about 1.3 million subscribers, reaching more than 1% of all U.S. households. The mean market penetration of all magazines in all regions is 1.53%. For the metro-level penetration data to be useful, market penetration for the same magazine/year should vary across regions. Table 11 in the Appendices provides a snapshot of pen-

¹The AAM is formerly known as the Audit Bureau of Circulation, or ABC. The name was changed in 2012. Fan (2013) and Gentzkow et al. (2014) use the analogous AAM reports on newspaper circulation; the coverage reports on magazines are relative new.

etration of all magazines. From that, one can see substantial geographical variation in a magazine's penetration levels for all magazines in 2012. Indeed, similar patterns are observed in all years.

For comparison, I include means of magazine attributes for the panel of 64 magazines. The sample of 34 magazines for my main analysis tends to have slightly more circulation on average, more ad pages, higher ad prices and less content pages. This reflects partially that they are the major players in the most popular genres.

In Table 3, I also include the summary statistics of two more variables constructed from the main variables. They are not used in the estimation, but are important for us to understand large magazine platforms. First, the mean per-page ad price is 15 cents per subscriber or \$150 per thousand household. The latter measurement is often called "cost per mille" (CPM) or cost per thousand consumer in the advertising industry. Magazines that I study here have very large CPMs, ranging from \$10 to \$1240. In other media like online search engines, newspapers, radio and TV, CPMs usually lie between \$5 to \$30, although CPMs in print media may be not directly comparable to TV and websites.² For example, the CPM for Super Bowl ads is between \$25 and \$30. It follows that advertisers must have derived large benefits per eyeball in these magazine platforms. Second, the average ad-to-non-ad-page ratio is approximately 1 : 1. This is consistent with the industry average (Magazine Publishing Association 2013). It is high in comparison to ads in other media. For instance, Wilbur (2008) reports a ratio between 1 : 5 to 1 : 4 in various television programs. In other words, consumers encounter more ads relative to the content in magazines than in other media.

²The advertising industry uses the term "CPM" indiscriminately in all media. However, as pointed out by Tirole and Rochet (2005), the CPM is charged based on *membership* in print media, while in TV, radio and websites, it is often based on usage.

I supplement the main data on magazine circulation and attributes with two additional sets of data. I use data on consumer rankings of U.S. magazines from the 2013 *Survey of the American Consumer*, administered by the media research firm GfK MRI. Along with demographics and other questions, the survey particularly asks consumers to rate magazines that they purchase as their best choice and so on. I observe percentages of each rating group for magazines in 2013. I interpret the data as proportions of consumers who regard a magazine as their first to fourth choices, corresponding to the functions in Section 4.1. These data allow me to construct more moments to identify demand parameters, especially the utility decrease parameter κ , which is usually unidentified with only market-level sales data. Finally, household demographics come from the American Community Survey micro-data available on IPUMS.

5.2 Descriptive Evidence

The model that I describe in Section 4 implies certain testable patterns in the data. First, the reader demand function (8) implies that magazines should have more consumers regarding them as the first best if they sell more subscriptions. This is the case because there is a one-to-one mapping between market penetration and consumers' mean product utilities. Since higher mean product utilities also lead to more first-best consumers, large platforms should have more first-best consumers. To investigate this relationship in the data, I run a regression of the number of first-best consumers of a magazine on its total subscription. I have only 34 product-level observations since I only have the survey data for the year of 2012-2013. I report the coefficient in Table 4. Despite the small sample size, the correlation is clearly present. In addition, the overall fit (i.e., $R^2 = 0.95$) seems surprisingly good given the sample size. In this regard, the pattern shown in the data is consistent with my model.

Table 5.1: Summary Statistics of Main Variables

	Main sample for analysis (34 magazines)				(64 magazines)
	Mean	SD	Min	Max	Mean
Market penetration (%)	1.53	1.21	0	27.5	-
Total circulation (1000)	1384.86	898.84	362.70	4209.68	1324.12
Subscription price (\$)	17.64	6.99	8.42	50.60	20.60
Ad pages (100 pages)	13.19	6.43	2.31	34.86	11.56
Ad price (\$1000/page)	166.24	120.10	46.71	1400.57	147.73
Content pages (100 pages)	13.17	3.49	7.21	25.22	14.27
Frequency (issues/year)	12.30	2.74	8	27	15.97
<i>Additional summary stats:</i>					
Per-subscriber ad price (\$)	0.15	0.12	0.01	1.24	0.14
Ad/total page ratio (%)	48.25	9.21	14.95	70.11	47.17

Notes: For market penetration, the unit of observation is a magazine/MSA/year. The number of observation is 110,419. For other product-level variables, the unit of observation is a magazine/year. The number of observation is 340 for 34 magazines in 10 years.

One important goal of this paper is to estimate advertiser demand when consumers multi-home. Specifically, a magazine's advertising price should reflect not only the number of its subscribers but also the number of consumers who can be reached only through its platform, *ceteris paribus*. Since the number of exclusive consumers on each platform is not observed in the data, running an OLS regression of ad prices on subscription levels without information on the reader composition would overestimate the weight of a consumer. However, equation (9) states that the number of exclusive consumers is a (non-linear) function of the number of first-best consumers. It suggests that the number of first-best consumers is correlated with the number of exclusive consumers. Table 5 presents the results when I use the number of first-choice consumers as an explanatory variable. In specification (1), I run an OLS of ad price on total subscription. The estimates says that, *ceteris paribus*, a magazine's ad price (per page) increases by \$72.32 for every 1,000 additional subscribers. In my model, it is equivalent to say that advertisers value consumers at \$.72 per eyeball. In specification (2), I include the number of first-best consumers as a proxy for unobserved exclusive eyeballs. The estimated coefficient on the number of subscribers decreases to 43.44. This is evidence that overlooking consumer composition when consumers multi-home generates bias on the advertiser demand estimates. Therefore, it is important to account for the value of exclusive consumers in the study of advertiser demand and platform pricing in two-sided markets.

Table 5.2: Larger Platforms with More First-Choice Consumers

Dependent variable: Number of 1 st -choice subscribers	
Total subscription	0.24*** (0.02)
R^2	0.95
No. magazines	34

Notes: Control variables include publisher dummies. The degree of freedom is 22.

Table 5.3: Using the Number of First-Choice Subscribers to Explain the Ad Price

Dependent variable: per-page ad price		
	(1)	(2)
Total subscription	72.32*** (6.70)	43.44*** (13.25)
No. 1st-choice subscribers	-	153.03*** (62.57)
R^2	0.89	0.92
No. of magazines	34	34

Notes: Subscription levels are measured in 1,000 subscribers. Control variables include the number of ad pages and a full set of publisher dummies. The degree of freedom is 20.

Chapter 6

Empirical Implementation

6.1 Econometric Specifications

In this section, I make further functional-form assumptions on the reader utility function (4.1) and the linear inverse advertising demand function (4.8) described in Chapter 4. The data used for estimation of the reader demand is a three-way (product-metropolitan-year) panel while the data used for estimation of the advertiser demand is a more conventional panel with magazine/product information in all years. Each of the reader utility and the advertiser demand function is specified to best suit the level of detailedness of available data. Discussion on identification of the parameters is presented in Section 6.2.

6.1.1 Specification of the Reader Demand Model

As in a standard random-coefficient logit model, I assume that the reader utility function (4.1) takes the following functional form:

$$u_{ijct} = \gamma_{ict}a_{jt} + \alpha_{ict}p_{jt}^s + x_{jt}\beta + \xi_{jt} + \Delta\xi_{jct} + \varepsilon_{ijct}; \quad (6.1)$$

where a_{jt} is the amount of advertising carried in magazine j in year t ; p_{jt} is the annual subscription price; x_{jt} is a vector of magazine characteristics, including the amount of non-advertising content pages, frequency of publication, and a (time-invariant) brand dummy. The parameter ξ_{jt} captures any change in the unobserved quality of magazine j in year t , while $\Delta\xi_{jct}$ captures region-time specific tastes for j and is assumed to be of mean zero. Notice that because the magazines of this study are the major national brand ones, consumers in different metropolitan areas face the same set of magazine characteristics at a time.

Like in Berry, Levinhson, and Pakes (1995), there are two types of heterogeneity in consumer preferences. First, recall that ε_{ijct} is the individual-specific taste shock and is i.i.d. with Type I extreme value distribution. Second, consumers have heterogeneous preferences for magazine prices and ad levels. I define the random coefficients as:

$$\begin{bmatrix} \gamma_{ict} \\ \alpha_{ict} \end{bmatrix} = \begin{bmatrix} \bar{\gamma} \\ \bar{\alpha} \end{bmatrix} + \begin{bmatrix} \gamma_1 \\ \alpha_1 \end{bmatrix}' z_{ict} + \Sigma\nu_{ict}, \quad (6.2)$$

where z_{ict} includes household demographics with $z_{ict} \sim P_{z_c}(z)$; ν_{ict} is the unobserved tastes attached to magazine subscription price and ad level, and $\nu_{ict} \sim P_\nu(\nu)$. $P_{z_c}(z)$ is the joint (empirical) distribution of demographics, and $P_\nu(\nu)$ is assumed to be standard normal. γ_1 and α_1 are vectors of parameters. The sum $\bar{\gamma} + \gamma_1' \bar{z}_{ict}$ - which represents the average consumer attitude towards advertising in media - is of particular interest since the current empirical literature on newspapers and magazines provides conflicting results on its sign.

Utility from the outside option, $j = 0$, is given by

$$u_{i0ct} = \delta_{i0ct} + \varphi T - \omega z_{ict} + \varepsilon_{i0ct}, \quad (6.3)$$

where δ_{i0ct} , mean utility of the outside option, is normalized to zero. T includes a

set of year dummies, and ε_{i0ct} is i.i.d. with Type I extreme value distribution. I allow demographics, z_{ict} , to affect the decision whether or not to subscribe to any magazine, in addition to interacting them with product characteristics in equation (6.1). I interpret the year dummies T as the effects of the Internet on traditional print media. In fact, T may be viewed as a "reduced-form" factor that reflects any change in the market environment over the years. For example, a lot of entries or exits can change T . However, as documented in Chapter 3, entry/exit does not raise concerns in the consumer magazine markets - especially for the major magazine brands in those major genres. *A priori*, the magnitude of T should increase over time as media consumers find it increasingly more attractive to use the Internet for information and entertainment. A similar interpretation is used in Fan (2013).

Using the notation introduced in Chapter 4, I can write equivalently:

$$u_{ijct} = \delta_{jct} + \vartheta_{ijct} + \varepsilon_{ijct}, \text{ for } j = 1, \dots, J; \text{ and } u_{i0ct} = \varepsilon_{i0ct}, \quad (6.4)$$

$$\text{where } \delta_{jct} = \bar{\gamma}a_{jt} + \bar{\alpha}p_{jt} + x_{jt}\beta - \varphi T + \xi_{jt} + \Delta\xi_{jct},$$

$$\vartheta_{ijct} = \omega z_{ict} + \gamma'_1 z_{ict} a_{jt} + \alpha'_1 z_{ict} p_{jt} + \gamma_2 \nu_{ict} a_{jt} + \alpha_2 \nu_{ict} p_{jt}.$$

Again, δ_{jct} represents the mean utility for magazine j in market ct , and together, the terms ϑ_{ijct} and ε_{ijct} capture household-specific-tastes deviating from the regional mean. Both ε_{ijct} and ν_{ict} are known to consumers but unobservable to firms and econometricians.

6.1.2 Specification of the Advertiser Demand Model

I specify the linear inverse demand function for advertising as following:

$$p_{jt}^a = \lambda_0 + \tilde{\lambda}_1 \tau_{jt}^e + \lambda_2 N_{jt}^s + \lambda_3 a_{jt} + \eta_j + \phi T + \epsilon_{jt}; \quad (6.5)$$

recall that p_{jt}^a is magazine j 's per-page advertising price, τ_{jt}^e the number exclusive readers on platform j , N_{jt}^s the total number of j 's subscribers, and a_{jt} the annual ad quantity.

Using the notation of equation (4.8), I assume that $y_{jt} = \eta_j + \phi T + \epsilon_{jt}$. In words, the total effect of demand shifters other than the total number of subscribers and the number of exclusive consumers, y_{jt} , includes a magazine fixed effect, η_j , a year fixed effect, T , and an i.i.d., zero-mean demand shock, ϵ_{jt} . Notice that in some earlier works - for example, Rysman (2004) and Fan (2013) - demographic information about regional media consumer markets are often included to explain variation in advertising prices of media across geographic markets. However, in this case, advertisers do not target consumers in a particular geographic market. Because all magazines are national-brand, advertisers face the same set of potential consumers for each magazine in a given genre. η_j removes any time-invariant factor that shifts j 's advertising demand while T captures any year-specific factor common to all magazines.

Wilbur (2008) uses a similar linear specification of advertiser demand in the U.S. television markets except that he does not include the number of exclusive consumers as a right-hand-side explanatory variable. A few other prominent studies of media markets, including Rysman (2004) and Fan (2013), estimate a log version of the advertiser demand equation. However, as shown in Rysman (2004), the log advertiser demand equation has a considerably different microeconomic foundation from the one I describe in Chapter 4. In particular, it is not consistent with inclusion of the number of exclusive consumers, which is the focus of this study.

6.2 Estimation of the Empirical Model

I estimate the reader demand equation (4.5) using the Method of Simulated Moments in spirit of both Berry, Levinsohn, and Pakes (1995), and Berry, Levinsohn,

and Pakes (2004). The parameters to be estimated are the consumer tastes for price, ad level, and magazine characteristics. In addition to the product-level moments originally used in BLP, I augment identification of the taste parameters with additional moments constructed from survey data on consumer ranking of magazines as described in Section ???. The advertiser demand equation (6.5) - and also (4.7) - is estimated separately using General Method of Moments (GMM). I estimate the two demand sides separately because of the large computational costs of the reader demand model and a data manipulation used to estimate the advertiser model. The same has been done in Berry and Waldfogel (1999), and Wilbur (2008). Berry and Waldfogel (1999) reports that separate estimation results are almost identical to joint estimation results.

Below, In Section 6.2.1, I first describe the moments used to construct moment conditions as well as the estimation procedure. In Section 6.2.2., I discuss issues regarding identification of the parameters and the instruments used to address potential endogeneity problems.

6.2.1 Moments

To estimate the reader demand, I use three sets of moments. First, I use the standard BLP moments that the unobserved product quality, ξ_{jt} , should be orthogonal to some instruments. Let ξ be a vector of unobserved product quality and \mathbf{Z}_1 a set of instruments. The first moment conditions are

$$G^1(\theta_1) \equiv E[\xi' \mathbf{Z}_1] = 0. \quad (6.6)$$

Since ξ_{jt} is not directly observed, I recover ξ_{jt} from the data. Specifically, let $\mathbf{s}(\delta(\theta))$ be a vector of predicted market penetration and \mathcal{S} be a vector of corresponding data. θ_1 consists of parameters entering into the (regional) mean utility part. I solve

for $\delta(\theta)$, which is the implicit solution to the system $\mathbf{s}(\delta) = \mathcal{S}$. In the case of a multiple discrete choice model, Fan (2013) establishes the invertibility of demand functions and proves the contraction mapping used in BLP is valid for computing δ . I extend her proof to my model, which details are included in Appendix A.2.

In step 1, I use simulation to approximate regional market penetration \mathbf{s} . For each metropolitan area and year, I randomly draw individuals from the census data, each characterized by their demographic characteristics and a population weight, ω . Conditional on the draws, I simulate unobserved taste parameter from the standard normal distribution, using antithetic acceleration to reduce variance introduced by simulation error, as suggested by Stern (1997). This step is done before the estimation process. The same draws are kept for uses in robustness checks and the counter-factual analysis.

For a guess of θ_2 , which are parameters entering into the individual specific tastes, the simulated market penetration is

$$s_{jct} = \sum_r^R \left(\Phi^1(\delta(\theta_1), \nu_i^r(\theta_2)) + \sum_{n=2}^4 \Phi^n(\delta(\theta_1), \nu_i^r(\theta_2)) \right) \omega_r; \quad (6.7)$$

where r denotes a simulated draw.

In step 2, I use the BLP contraction mapping to obtain a vector of product-region-year specific mean utility parameters δ_{jct} . Conceptually, the BLP contraction mapping converts a non-linear search problem (i.e., searching for δ such that $\mathbf{s}(\delta(\theta_1)) = \mathcal{S}$) to a linear one. So even δ has a large number of elements, the solution emerges quickly in practice. To recover ξ_{jt} with known δ , I define $\tilde{\xi}_{jt} = \bar{\gamma}a_{jt} + \bar{\alpha}p_{jt} + x_{jt}\beta + \xi_{jt}$ as the mean utility for product j , including both mean tastes for observed characteristics and unobserved quality. It follows that $\delta_{jct} = \tilde{\xi}_{jt} - \varphi T + \Delta\xi_{jct}$. In step 3, as suggested by Nevo (2001), $\tilde{\xi}_{jt}$ is estimated by running a GLS regression of δ_{jct} on a full set of product dummies, in addition to the time dummies. Finally, we have

$$\xi_{jt} = \tilde{\xi}_{jt} - \bar{\gamma}a_{jt} - \bar{\alpha}p_{jt} - x_{jt}\beta.$$

Notice that the first set of moments are at the product level. To estimate θ_2 , I use the condition $\mathbf{s}(\delta(\theta)) = \mathcal{S}$ again. For given $\delta(\theta_1)$, I construct additional sample moments as follows:

$$G^2(\theta_2) \equiv \sum_t \sum_j \sum_c (\mathcal{S}_{jct} - s_{jct}(\theta_2)) \mathbf{Z}_2^r = 0, \quad (6.8)$$

where the vector of instruments \mathbf{Z}_2^r consists of average demographics interacting with exogenous product-level instruments. I use these moments to search for θ_2 with given δ .

The third set of moments matches the proportions of first to fourth choice consumers predicted by the model to the observed percentages from the consumer survey data. Let Φ_j be a vector with elements Φ_j^1 to Φ_j^4 as in (6). I stack Φ_j for all j to construct the vector Φ . Let Φ^{data} be a vector of corresponding data. The third set of moments is then given by:

$$G^3(\theta) \equiv (\Phi^{data} - \Phi)' \mathbf{Z}_1 = 0. \quad (6.9)$$

As I discuss in the next section, the third set of moments help identify the utility loss parameter κ .

I stack the moments and search for θ that minimizes the weighted distance. Formally,

$$\theta^* = \arg \min_{\theta} G(\theta)' \mathcal{W} G(\theta), \quad (6.10)$$

where

$$G(\theta) = \begin{pmatrix} G^1(\theta_1) \\ G^2(\theta_2) \\ G^3(\theta) \end{pmatrix},$$

and \mathcal{W} is a positive definite weighting matrix. I follow the standard method by first assuming homoscedasticity and using $(Z'Z)^{-1}$ to obtain a consistent estimate of θ . I then use this estimate to get $EG(\hat{\theta})G(\hat{\theta})'$, which is subsequently used to get the final estimate of θ .

The advertiser demand (6.5) is estimated separately from the reader demand side using GMM. I use estimates from the reader demand to predict the number of exclusive eyeballs on each platform, which enters equation (6.5) as an explanatory variable. The moment condition is that the residual term ϵ_{jt} is orthogonal to some instruments \mathbf{Z}_3 .

6.2.2 Identification

In this study, there are two issues of identification. The first issue is about identification of non-linear models. In general, besides any potential endogeneity problem, covariation of market penetration and relevant right-hand-side product attributes identifies parameters in θ_1 that enters into the mean utility part in (4.5). Geographical variation in penetration of a same magazine and variation in demographic information across metropolitan area identify parameters in θ_2 that enters into the individual specific tastes. The utility loss parameter κ is identified with the survey data. In the literature, parameters with a similar connotation are identified either with individual-level data (e.g., Gentzkow 2007) or under the assumption that some markets are monopolies (e.g., Fan 2013). The survey data provide another unique opportunity to identify κ . Conceptually, the percentages of first choice consumers Φ_j^1 are independent of κ while others are functions of κ . Suppose I can estimate θ

solely from Φ_j^1 , which are similar to “market share” function generated by a standard mixed logit model. I can then calculate percentages of second-choice to fourth-choice consumers based on the estimates while assuming $\kappa = 0$. The difference between the calculated percentages and the data is necessarily explained by κ .

The second issue of identification concerns potential endogeneity of some of the variables. In the estimation of reader demand, the subscription price p_{jt}^s and the number of advertising page a_{jt} may be endogenous since they can correlated with the unobserved quality component ξ_{jt} . Due to the three-way panel structure of the data, I can include brand dummies and time dummies in the estimation. These dummy variables remove any unobserved product-specific and time-specific factors. To further address the endogeneity problem, I use three sets of instrument variables (IVs) commonly used in the empirical IO literature. The first set of IVs consists of “BLP-type instruments”. They include the number of products and average characteristics of other products of the same genre. In principle, they should satisfy instrument relevance because of the oligopoly market structure. They are exogenous because it is assumed that product characteristics are pre-determined. The second set of so-called “Hausman-type instruments” includes the average subscription price, ad page number, and content page number of all other products that belong to different genres but the same publisher. They are correlated with own price and ad pages due to cost-side factors that are common to a publisher. They are exogenous because each genre is a separate market segment. Kaiser and Song (2009) use similar instruments in the context of German magazines. I also include a full set of publisher dummies to account for any unobserved time-invariant cost factors. After controlling for product fixed effects, publisher dummies should be uncorrelated with unobserved consumer tastes.

In MSM or GMM, there is no “first stage” as in two stage least stage (2SLS). In order to check validity of the instruments, I run regressions similar to the first-stage

Table 6.1: Instrument Relevance for Endogenous Price and Ad pages

	Endogenous variables	
	Subscription price	Ad pages
<i>Included instruments</i>		
Content pages	1.36*** (0.09)	1.29*** (0.10)
Frequency of publication	-0.60*** (0.20)	0.20 (0.23)
<i>Excluded instruments</i>		
IV BLP 1	-0.28* (0.15)	-0.10 (0.17)
IV BLP 2	-0.62*** (0.13)	0.09 (0.15)
IV BLP 3	-0.58*** (0.21)	0.25 (0.25)
IV Hausman 1	0.14 (0.39)	-0.82* (0.46)
IV Hausman 2	7.92*** (1.24)	0.42** (0.18)
IV Hausman 3	-0.05 (0.15)	-0.76 (1.34)
Publisher dummy 1	-17.93*** (3.43)	-10.27** (4.91)
Publisher dummy 2	-13.42*** (4.02)	-0.88 (1.41)
Publisher dummy 3	-17.13*** (4.15)	-1.30 (1.03)
Publisher dummy 4	-15.81*** 4.81	1.49 (1.20)
Publisher dummy 5	-5.43* (2.91)	-16.15*** (4.04)
Publisher dummy 6	2.81 (2.78)	-8.66* (4.44)
Publisher dummy 7	-1.88 (2.67)	-9.76** (4.53)
Publisher dummy 8	-6.45** (2.87)	-4.76 (4.62)
Publisher dummy 9	16.31*** (3.83)	-5.81 (4.93)
Publisher dummy 10	-14.70*** (4.37)	-6.10 (5.14)
R^2	0.79	0.66
F-test	59.08	29.96

Notes: This table summarizes estimates from “first stage” regressions on the endogenous variables. The degree of freedom is 290. *, **, *** denotes 1%-, 5%-, and 10%- significance level, respectively.

regression in 2SLS by regressing the endogenous variables on both the included and excluded instruments. I present the “first stage” results in Table 6.1. First, I verify whether they are relevant instruments. In the subscription price equation, coefficients on excluded instruments are mostly significant, and their signs seem largely intuitive. In the ad page equation, many cost-related instruments, such as average subscription price and ad pages of other genres by the same publisher (i.e., IV Hausman 1 and IV Hausman 2) and publisher dummies, are highly correlated with ad pages. Second, I see whether there is a problem of weak instruments. From the first-stage regressions, I report the F-statistics that are sufficiently larger than the critical values for the weak IV test. So jointly, the instruments pass the test for weak IVs. For additional robustness checks, for each endogenous variable, I run three separate “first stage” regressions, removing one set of instruments at a time. For estimates from the “first stage” regressions using subsets of IVs, see Table 6.2. Then, I check correlations among the predicted values of an endogenous variable: for the final estimates to be robust to the choice of instruments, the predicted values using only parts but not all of the instruments should be highly correlated. Table 6.3 summarizes their correlations. For example, the correlation between the predicted value of subscription price without the first set of IVs, \hat{p}_1^s , and the predicted value without the second set of IVs, \hat{p}_2^s , is 0.92. The correlation between the predicted value of ad pages without the first set of IVs, \hat{a}_1 , and the predicted value without the second set of IVs, \hat{a}_2 , is 0.99. Because the predicted values are highly correlated, the results are not sensitive to the inclusion of any particular set of instruments.

For the advertiser demand, I use similar instruments to address the endogeneity of ad pages, the subscription level and the number of exclusive eyeballs. However, one may argue the endogeneity problem in this case is somewhat questionable since advertisers often observe a similar amount of information when they purchase ad spots on magazines. Therefore, there should not be any omitted variable left in ϵ_{jt} ,

which is then a pure demand shock to advertisers.

Table 6.2: Robustness Checks on the instruments

Endogenous variable:	Subscription price			Ad page		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Included IVs</i>						
Constant	-62.36*** (17.43)	20.42*** (3.14)	1.94 (2.32)	19.44 16.70	8.52*** (3.37)	-16.12*** (2.28)
Content pages	0.81*** (0.08)	1.38*** (0.09)	1.47*** (0.10)	1.32*** (0.08)	1.31*** (0.10)	1.12*** (0.10)
Frequency of publication	0.06 (0.22)	-0.52*** (0.21)	0.37*** (0.12)	0.15 (0.22)	0.17 0.22	0.51*** (0.12)
<i>Excluded IVs</i>						
IV BLP 1	-	-0.60*** (0.14)	-0.55*** (0.14)	-	-0.04 (0.15)	-0.14 (0.13)
IV BLP 2	-	-0.23* (0.13)	-0.37*** (0.11)	-	0.11 (0.14)	0.31*** (0.11)
IV BLP 3	-	-0.42** (0.19)	0.04 (0.15)	-	0.18 (0.21)	0.25* (0.15)
IV Hausman 1	0.33* (0.17)	-	-0.26** (0.13)	0.41** (0.17)	-	0.15 (0.13)
IV Hausman 2	-0.67 (0.51)	-	0.05 (0.18)	-0.43 (0.49)	-	0.10 (0.17)
IV Hausman 3	3.77*** (0.97)	-	0.13* (0.08)	1.25 (0.92)	-	0.04 (0.08)
Publisher dummies	Yes	Yes	No	Yes	Yes	No
(Adjusted) R-squared	0.67	0.73	0.64	0.65	0.64	0.59
F-stat	45.67	59.96	74.65	41.48	40.07	62.41

Notes: This table summarizes estimates from "first stage" regressions on the endogenous variables using some but not all of the instruments. The degree of freedom is 290. *, **, *** denotes 10%, 5%, and 10%- significance level, respectively.

Table 6.3: IV Robustness

Predicted subscription price				
	\hat{p}^s	\hat{p}_1^s	\hat{p}_2^s	\hat{p}_3^s
\hat{p}^s	1			
\hat{p}_1^s	0.94	1		
\hat{p}_2^s	0.98	0.92	1	
\hat{p}_3^s	0.91	0.85	0.94	1

Predicted ad pages				
	\hat{a}	\hat{a}_1	\hat{a}_2	\hat{a}_3
\hat{a}	1			
\hat{a}_1	0.99	1		
\hat{a}_2	0.99	0.99	1	
\hat{a}_3	0.95	0.95	0.96	1

Notes: \hat{p}^s is the predicted value of subscription price using all 3 sets of IVs, and \hat{p}_1^s is the predicted value of subscription price using all but the first set of IVs, etc.

Chapter 7

Estimation Results and the Counterfactual Analysis

7.1 Estimation Results

7.1.1 Reader Demand Estimates

In this section, I present results of reader demand estimation. Table 7.1 summarizes the estimates and standard errors of the main coefficients. First, the estimate of the baseline price coefficient, $\bar{\alpha}$, and the implied price coefficient for an average consumer are both significantly negative as expected. Recall that, for subscription of an average magazine, consumers only pay a little more than one dollar per issue. This insures that consumers experience disutility from paying a price and that the reader demand curve is downward sloping even when prices are relatively small for magazine subscription. Second, the estimate of the baseline advertising coefficient, $\bar{\gamma}$, and the implied advertising coefficient for an average consumer are also significantly negative. After addressing the endogeneity issue, I interpret these estimates as evidence of consumer ad-aversion; that is, average consumers dislike pages devoted to advertising in a magazine. Together with the estimated price coefficient, this implies an ad nuisance

cost at about five cents: for average consumers, one more page of ads is equivalent to an increase in the annual subscription price by \$0.05. It further implies that the full price - subscription price plus ad nuisance costs - that average consumers pay for an average magazine annual subscription is approximately \$83, in which the subscription fee only accounts for a bit more than 25%, and ad nuisance costs account for the rest 75%. In other words, consumers pay more in terms of ad nuisance costs than they nominally pay to purchase an annual subscription. Third, there is some significant heterogeneity in consumer tastes for prices and advertising. Here, in practice, I only maintain the estimates for interactions between product characteristics and individual demographics that are significantly different from zero. For example, the estimate of the interaction term between price and (log) household annual income, α_1 , suggests that higher-income consumers are less sensitive to a change in magazine subscription. Similarly, the estimates suggest that households with a higher male ratio dislike magazine advertising, and that consumers with less school years and lower income dislike ads more. Although I do not include a model of advertising technology - that is, how advertising messages translate to consumer purchases or sales, I suggest that these demographic-specific ad taste parameters may be interpreted as effectiveness of advertising messages to different consumer types. Overall, my results suggest that male households with lower income and less school years find magazine advertising less appreciable and thus less useful. A caveat is that consumers differ only slightly in their attitudes towards magazine advertising given the magnitude of coefficients of those interaction terms.

In summary, my main results regarding consumer ad-aversion is in contract to the findings in other empirical studies on print media. For example, various works find that readers are ad-neutral in newspaper markets.¹ Using data on German magazines, recent works find ad-loving consumers. In particular, Kaiser and Song (2009) find

¹See Gentzkow (2007) and Fan (2013) for examples with structural models. A detailed survey on related studies is provided by Chandra and Kaiser (2016).

that, in six genres of German magazines, consumers prefer magazines that have higher ad ratios. Their results can be puzzling since magazines should not want to include any non-ad content pages if readers strictly prefer ads to non-ad pages. My finding is largely consistent with findings in other media, such as in television and radio.

Quantifying and understanding cross-group externalities and indirect network effects is important in two-sided markets. To investigate issues behind the identification of consumer ad preference parameter, I estimate a series of logit equations with different specifications. The results are reported in Table 7.2. Specification (1) is logit model without year fixed effects and product fixed effects, (2) is logit with year and product fixed effects, (3) and (4) are the IV versions of those with instruments as described in Section 5. In my case, the negative correlation between the number of advertising pages and magazine penetration is present in the original data. This is reflected in specification (1), in which the coefficient on ad pages is negative when no effect has been taken to address the endogeneity problem of ad volume. However, one can see both subscription price and ad pages suffer from severe endogeneity issues. This is seen by comparing specification (1) and (3), and (2) and (4). When I use instruments, the coefficient on price and advertising all become more negative, which confirms our suspicion that magazines have higher unobserved quality charge higher prices and have more advertising pages. Nevertheless, I cannot rule out the possibility that difference between my findings and other empirical works is due to cross-country and/or cross media differences in consumers' attitude towards advertising. For example, Kaiser and Song (2009) have product fixed effects and instruments similar to ones used in this study. Yet, they find the opposite results with German magazine data.

By comparing specification (4) in Table 7.3 to my main estimates, I observe that the IV logit model suffers from model specification problems. In particular, the logit model overestimates the coefficients on content pages - a product attributes that

 Table 7.1: Estimates for Main Reader Demand Parameters

Variable	Parameter	Estimate	Standard Error
Subscription price	$\bar{\alpha}$	-0.267	0.015
<i>Interacting with</i> log(HH income)	α_1	0.022	0.001
Ad pages	$\bar{\gamma}$	-1.039	0.203
<i>Interacting with</i> log(HH income)	γ_{11}	1.1E-4	1.8E-5
HH sex ratio	γ_{12}	-3.6E-4	1.1E-4
education	γ_{13}	1.2E-5	5.60E-06
Content pages	β_1	0.031	0.018
Frequency	β_2	0.022	0.003
Diminishing utility	κ	-1.920	0.387
	φ_{2004}	-0.009	0.006
	φ_{2005}	-0.017	0.006
	φ_{2006}	-0.004	0.006
	φ_{2007}	-0.006	0.007
Year dummies	φ_{2008}	-0.039	0.008
	φ_{2009}	-0.028	0.008
	φ_{2010}	-0.074	0.008
	φ_{2011}	-0.048	0.008
	φ_{2012}	-0.036	0.008

consumers like, and underestimates consumers' ad-aversion. In the logit model, one more content page in a year translates to a decrease of subscription price by 1 cent. However, in the main model, the equivalent decrease is approximately 0.1 cent, with other product attributes and quality remained constant. The coefficient on advertising also nearly doubles from 0.57 to 1.03. In summary, incorporating consumer multi-homing behavior into the demand model is crucial since doing that would correct for the model specification error when single-homing is assumed.

Table 7.1 also reports the estimated coefficients on time dummies. As discussed in Section 3, time dummies capture decreases in demand for magazines due to the rise of the Internet. All coefficients are negative and highly significant. For example, the decrease in demand in the year 2010 is equivalent to an increase of subscription price by \$3.5, or about %20 for an average magazine. The drop in demand due to the Internet is indeed very important.

Table 7.3 indicates that consumer heterogeneity plays an important role in tastes for magazines, especially in determining whether or not to purchase magazines at all. In addition, the role of consumer heterogeneity is qualitatively and sometimes qualitatively different across genres. In the full model, I allow genre dummies to interact with a few parameters. Most interestingly, I find that consumers' ad preferences vary only slightly across genres. For example, the results suggest that consumers are less averse to advertising in men's magazines and women's magazines with a general focus. However, magnitudes of such differences are almost negligible. On the other hand, consumer heterogeneity affects consumer decision to purchase different genres of magazines. In general, consumers with fewer years of education tend to buy magazines less often. consumers who have recently moved from another states and hence are less "settled", tend to purchase magazine subscriptions less often. Consumers who live in households with more male members are more likely to buy men's magazines and personal finance magazines while those live with more female members tend to buy

more interior design (or “shelter”) magazines besides all kinds of women’s magazine.

Table 7.2: Inclusion of Brand Dummies and Instruments

Dependent variable: $\ln(s_{jct}) - \ln(s_{oct})$				
	Logit		IV logit	
	(1)	(2)	(3)	(4)
Price	-0.046*** (0.000)	-0.001 (0.001)	-0.088*** (0.001)	-0.102*** (0.006)
Ad pages	-0.344*** (0.007)	-0.024*** (0.001)	-0.897*** (0.000)	-0.566*** (0.063)
Content pages	0.843*** (0.014)	0.022 (0.015)	1.881*** (0.020)	1.088*** (0.105)
Frequency	0.067*** (0.001)	0.022*** (0.002)	0.124*** (0.001)	0.015** (0.006)
Product dummies	No	Yes	No	Yes
Year dummies	No	Yes	Yes	Yes
Ad nuisance cost	-0.7 cents	-	-1.0 cents	-0.6 cents
Similar dummies and/or IVs used in:	-	Kaiser and Wright (2006, IJIO)	Fan (2013, AER)	Kaiser and Song (2009, IJIO)
No. of observations	110419	110419	110419	110419

Notes: Kaiser and Wright (2006), and Kaiser and Song (2009) report positive coefficients for ad level. Kaiser and Song (2009) uses ad/content ratio as the measurement of ad level. Fan (2013) reports a close-to-zero, non-significant coefficient for ad level in footnote 8; thus, in her main specification, it is assumed that readers are ad-neutral.

Table 7.3: Genre-specific parameters

Variable	Genres		
	<i>Women's health</i>	<i>Shelter</i>	<i>Women's general</i>
Ad pages	(Baseline)	-4.8E-5 (2.7E-5)	7.8E-5 (2.8E-5)
log(income)	-0.176 (0.036)	0.061 (0.039)	-0.499 (0.035)
Age	0.001 (0.001)	0.020 (0.002)	0.012 (0.002)
HH sex ratio	-0.812 (0.206)	-0.915 (0.208)	-0.930 (0.212)
Education	0.244 (0.013)	0.186 (0.013)	0.251 (0.013)
Home ownership	-0.474 (0.126)	0.427 (0.067)	-1.242 (0.055)
Migration	-0.645 (0.126)	-0.513 (0.125)	-0.365 (0.103)
	<i>Men's</i>	<i>Women's fashion</i>	<i>Personal finance</i>
Ad pages	1.4E-4 (3.3E-5)	-4.4E-5 (2.5E-5)	4.2E-5 (2.6E-5)
log(income)	-0.394 (0.041)	0.210 (0.034)	-0.217 (0.047)
Age	-0.003 (0.002)	0.005 (0.001)	0.009 (0.002)
HH sex ratio	2.654 (0.223)	-0.914 (0.210)	0.608 (0.247)
Education	0.247 (0.015)	0.115 (0.012)	0.204 (0.016)
House ownership	0.592 (0.079)	1.298 (0.040)	0.254 (0.078)
Migration	-0.785 (0.147)	-0.846 (0.075)	-0.584 (0.145)

7.1.2 Advertiser Demand Estimates

In this section, I present results of advertiser demand estimation. From the reader demand model, I predict the numbers of exclusive eyeballs on each platform, which information is missing in the data. I find, on average, 20% of a magazine's subscribers are exclusive to the platforms. I use the predicted values as an explanatory variable in the inverse advertiser demand function. I report the estimates and standard errors of advertiser demand parameters in column (1) of Table 10. I find that an increase in subscription level by 1,000 subscribers would lead to an increase in advertising price by \$49, holding other factors constant. Moreover, an increase in the number of exclusive subscribers by 1,000 would raise the advertising price by \$71. In context of the advertiser model in Section 3, the results imply that advertisers value exclusive readers at \$0.12 per eyeball while they value non-exclusive readers at \$0.05 per eyeball. In other words, they value exclusive readers twice as much as they value non-exclusive ones. This is direct evidence that platform advertising price reflects advertiser differential valuation of single-homing and multi-homing consumers, which is hypothesized in the recent theoretical literature on two-sided markets.

As discussed in Section 5, advertising pages, subscription levels and number of exclusive eyeballs are endogenous. To investigate the effects of endogeneity on the estimates, I run three additional regressions which results are also reported in Table 10. In specification (1), I run an OLS of the advertising price on ad pages and the subscription level while controlling for product fixed effects and year fixed effects. I repeat the same regression with the larger panel of 64 magazines in (2). Comparing these, I find coefficients are almost the same with both samples. I contend that my main results on the advertiser side should be externally valid. In specification (3), I include the predicted numbers of exclusive eyeballs in a similar OLS regression. The result suggests that an increase in the number of exclusive consumers would lead to \$0.41 increase in the advertising price. The estimated coefficient on subscription level

Table 7.4: Estimates for Advertiser Demand

Variable	Parameter	Estimate			
		Main	(1)	(2)	(3)
Ad pages	λ_3	-17.93 (8.90)	-105.52 (11.92)	-105.53 (8.34)	-111.9 (12.23)
No. subscribers	λ_2	49.42 (19.89)	52.36 (8.47)	46.42 (5.55)	1.60 (20.99)
No. exclusive eyeballs	$\tilde{\lambda}_1$	71.11 (15.92)	-	-	414.22 (156.87)
Product dummies	η_j	Yes	Yes	Yes	Yes
Year dummies	ϕ	Yes	Yes	Yes	Yes
No. observations		340	340	640	340

is small and insignificant. In other words, once I include the number of exclusive eyeballs which is omitted in specification (1) and (2), the result suggests that advertisers value exclusive eyeballs at \$0.41 per eyeball while they do not value non-exclusive eyeballs at all. In context of our model, it implies that consumers' attention beyond first impressions is worthless to advertisers. However, once the endogeneity issue is addressed, the estimates change and lead to more reasonable interpretations.

7.2 Counterfactual Exercise: The Impact of the Internet

In this counterfactual exercise, I simulate the market outcomes in 2012 if the reader demand for magazines were as strong as in 2003.² I interpret the counterfactual results as the effects of the Internet on magazine subscription markets and advertising markets. From Table 5, I estimate a set of time dummies in consumers' indirect utility function. As discussed in Section 3, those dummies would capture any time-specific change in the value of the outside option. Given that there is no entry/exit of major magazines in genres of this study, I attribute those time effects to increasing attractiveness of the Internet. Holding all other exogenous variables constant, I let t_{2012} to be zero, and simulate market outcomes based on demand and cost-side estimates from the estimation. Table 9 summarizes the results.

Several interesting observations emerge from Table 9. In comparison to what really happened in 2012, if the reader demand for magazines were as strong as a decade ago, the average subscription price of the 34 major magazines would have been \$3, or about 20% higher. Even prices are higher, now because of stronger demand, more magazine subscriptions are sold: the average subscription increases

²This counterfactual exercise is based on a simplified version of the model where the effect of ad volume on consumer utility is set to zero.

by 135,000 or 9%. The average number of single-homing consumers also increases by 70,000 or 22%. Together, they imply (net) market expansion for the six genres. In other words, not only some of the existing consumers buy more magazines, but new consumers also enter the market. With higher subscription levels and more exclusive consumers, strong reader demand translates to boosting effects in magazine advertising markets. Comparing to real data in 2012, platforms charge higher prices and admit more advertisers on average in the hypothetical case. As a result, the average ad revenue increases by \$72 million. Overall, I find that the Internet has very large effects on magazines. The direct impact of the Internet on the subscription market is considerably large while the indirect impact through reduced consumers and less favorable reader composition on advertising markets is even bigger - especially when magazines rely heavily on advertising revenues.

One caveat is that the above analysis is based on *all other factors being constant*. From 2003 to 2012, there are other changing factors besides the time-specific effects. For example, as I evaluate the counterfactual case, all other factors, such as consumer demographics, exogenous product attributes and unobserved quality, are assumed to take their 2012 value. This explains why outcomes in 2003 data are very different from the counterfactual outcomes in 2012 even strength of the reader demand is exactly the same.

Table 7.5: Counterfactual Results

Market Outcome (Mean)	2003 Data	2012 Data	2012 Counterfactual
p^s (\$)	22.04	15.14	18.63
N^s	1311.87	1441.22	1576.34
τ^e	324.03	312.22	383.05
a	14.27	11.24	14.64
p^a	131.75	166.47	174.58
<i>Additional outcome variables (in million \$):</i>			
Ad revenue	190	190	262
Circ. revenue	36.4	23.8	30.1

Chapter 8

Conclusion

Media platforms compete for both consumers and advertisers, especially when consumers divide their attention among multiple platforms. While traditional economics models assume consumers patronize a single platform, in this thesis, I model consumer demand for multiple magazines (“multi-homing”), and magazines subscription price and ad price decisions. Using a novel data-set on metropolitan-level magazine sales, characteristics and consumer ranking of magazines, I estimate the model and quantify the indirect network effects in magazine markets. I provide the first direct evidence that media ad prices reflect advertisers differential valuation of exclusive and non-exclusive eyeballs on platforms.

On the consumer side, I estimate a multiple discrete choice model of demand using new panel data on US magazine regional sales and characteristics from 2003 to 2012 and survey data on consumer rankings of magazines. Consumers have preferences over prices, ad volume, and other characteristics, and they have diminishing marginal utility from multiple purchases. Demand side results suggest that consumers ad nuisance cost is approximately 5 cents per ad page, in contrast to the ad-neutrality or ad-loving findings in the print media literature.

On the platform side, my model relates to the emerging theoretical two-sided

market literature that emphasizes the importance of multi-homing. In the model, magazines compete to catch more eyeballs and the accompanying advertising revenues. When consumers are reached on multiple platforms, exclusive eyeballs are more valuable to advertisers and platforms. I estimate that, on average, exclusive eyeballs value 7 cents more or twice as values of shared eyeballs, thus confirming predictions in the theoretical literature.

I use the estimation results to investigate how the market would have differed if demand for magazines remained as strong as in 2003. I interpret the results as possible effects of the Internet on magazine markets over the decade. Subscription prices would increase by 20% on average, exclusive readerships 22% higher, and therefore, ad prices would also increase. A model without consumer multi-homing tends to overestimate the value of a subscriber, but underestimates the power market of platforms.

One limitation of this analysis is that much of the empirical work - including the estimation and inferences - is based on consumer multi-homing data at the *aggregate* level. In particular, it is only known to us the proportions of first-choice to fourth choice consumers of each magazine, and such magazine-level data are only available for one year (i.e., year of 2012). This lack of individual-level data leads to my use of imputed data - instead of actual observed data - for the analysis of exclusive consumers on platforms, which is a key subject of this thesis. At the meantime, media platforms that have long histories of operation and constantly make pricing decisions probably do have access to individual-level data on consumer multi-homing and base their decisions on these data. However, I view relevant parts of my analysis as an invaluable first-step to understand the pricing mechanism in not only U.S. magazine markets specifically but also in general media markets, especially given the scarcity of consumer multi-homing data and the complexity of modeling such behavior. In fact, most of the empirical analysis - including the empirical model and the estimation procedure - can be easily extended to handle finer data on consumer

multi-homing in media markets. Therefore, future research should take advantage of individual-level data on consumer multi-homing, which are known to exist and be commercially available.

Bibliography

Ambrus, A., E. Calvano and M. Reisinger (2016), “Either or Both Competition: A Two-sided Theory of Advertising with Overlapping Viewerships.” *American Economic Journal: Microeconomics*, forthcoming.

Armstrong, M. (2006), “Competition in Two-Sided Markets”. *RAND Journal of Economics*, 37: 668-691.

Anderson, S. P. and S. Coate (2006), “Market Provision of Broadcasting: A Welfare Analysis.” *Review of Economic Studies*, 72 (4): 947-972.

Anderson, S. P., A. de Palma, J.F. Thisse (1992). *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA.

Anderson, S. P., Ø. Foros and H. J. Kind (2016), “Competition for Advertisers and for Viewers in Media Markets.” *Economic Journal*, forthcoming.

Anderson, S. P., Ø. Foros, H. J. Kind, and M. Peitz (2012), “Media market concentration, advertising levels, and ad prices.” *International Journal of Industrial Organization*, 30 (3), 321-325.

Anderson, S. P., and B. Jullien (2016), ”The Advertising-Financed Business Model in Two-Sided Media Markets .” in *The Handbook of Media Economics*, Simon Anderson, David Stromberg and Joel Waldfogel eds., Volume 1A, Elsevier.

- Anderson, S. P. and M. Peinz (2015), "Media Seesaws: Winners and Losers in Platform Markets." Working paper.
- Argentesi, E. and Filistrucchi, L. (2007), "Estimating Market Power in a Two-Sided Market: The Case of Newspapers." *Journal of Applied Econometrics*, 22: 1247-1266.
- Athey, S., E. Calvano and J. S. Gans (2014), "The Impact of the Internet on Advertising Markets for News Media." Working Paper, Harvard University.
- Augereau, A., Greenstein, S. and Rysman, M. (2006). "Coordination versus Differentiation in a Standards War: The Adoption of 56K Modems." *RAND Journal of Economics*, 37: 887-909.
- Beebe, J.H. (1977). "Institutional structure and program choices in television markets." *Quarterly Journal of Economics*, 91(1), 15-37.
- Berry, Steven (1994), "Estimating discrete-choice models of product differentiation." *RAND Journal of Economics*, 25: 242-262.
- Berry, S., J. Levinsohn, and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4): 841-890.
- Berry, S., J. Levinsohn, and A. Pakes (2004), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112 (1).
- Berry, S. and J. Waldfogel (1999), "Free entry and social inefficiency in radio broadcasting." *RAND Journal of Economics*, 30: 397-420.
- Caillaud, B, and B. Jullien (2003), "Chicken & egg: competition among intermediation service providers." *Rand Journal of Economics*, 34(2), 309-328.

Chandra, A. and A. Collard-Wexler (2009), "Mergers in Two-Sided Markets: An Application to the Canadian Newspaper Industry," *Journal of Economics and Management Strategy*, 18(4): 1045-1070.

Chandra, A. and U. Kaiser (2016), "Newspapers and Magazines." in *The Handbook of Media Economics*, Simon Anderson, David Stromberg and Joel Waldfogel eds., Volume 1A, Elsevier.

Depken, C.A. (2004), "Audience characteristics and the price of advertising in a circulation industry: evidence from US magazines." *Information Economy Policy*, 16, 179-196.

Depken, C.A. and D.P. Wilson (2004), "Is advertising a good or bad? Evidence from U.S. magazine subscriptions." *Journal of Business*, 77(2), S61-S80.

Fan, Y. (2013), "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market", *American Economic Review*, 103(5): 1598-1628.

Gentzkow, M. (2007), "Valuing New Goods in a Model with Complementarities: Online Newspapers," *American Economic Review*, 97 (3): 713-744.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson. (2014), "Competition and Ideological Diversity: Historical Evidence from US Newspapers." *American Economic Review*, 104(10).

Hendel, I. (1999), "Multiple-Discrete Choice Models: An Application to Computerization Returns." *Review of Economic Studies*, 66(2), 423-446.

Hiedemann B, M. Sovinsky, and S. Stern. (2013), "Will You Still Want Me Tomorrow? The Dynamics of Families' Long-Term Care Arrangements." Working Paper.

Hong, S.-H. (2007), "The Recent Growth of the Internet and Changes in Household-Level Demand for Entertainment," *Information Economics and Policy*, 19, 304-318

Jeziorski, P. (2014), "Effects of Mergers in Two-Sided Markets: The US Radio Industry." *American Economic Journal: Microeconomics*, 6(4): 35-73.

Kaiser, U. and M. Song (2009), "Do Media Consumers Really Dislike Advertising? An Empirical Assessment of the Role of Advertising in Print Media Markets." *International Journal of Industrial Organization*, (27): 292-301.

Kaiser, U. and J. Wright (2006), "Price Structure in Two-Sided Markets: Evidence from the Magazine Industry." *International Journal of Industrial Organization*, (24): 1-28.

Koschat, M. A. and W. A. Putsis (2000a), "Who wants you when you're old and poor? Exploring the economics of media pricing." *Journal of Media Economics*, 13(4), 215-232.

Koschat, M. A. and W. A. Putsis (2000b), "Audience characteristics and bundling: a hedonic analysis of magazine advertising rates." *Journal of Marketing Research*, 39(2), 262-273.

Nevo, A. (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry." *Econometrica*, 69(2), 307-342.

Oster, S. and F. Scott Morton (2005), "Behavioral biases meet the market: the case of magazine subscription prices." *Advanced Economic Analysis & Policy*, 5(1) Article 1.

Rochet, J.-C. and J. Tirole. (2003), "Platform competition in two-sided markets." *Journal of European Economic Association*, 1: 990-1029.

Rochet, J.-C. and J. Tirole. (2006), "Two-Sided Markets: A Progress Report." *RAND Journal of Economics*, 37 (3): 645-67.

- Rysman, M. (2004), "Competition Between Networks: a Study of the Market for Yellow Pages," *Review of Economic Studies*, 71 (2): 483-512.
- Rysman, M. (2009), "The economics of two-sided markets." *Journal of Economic Perspective*, 23(3): 125-143.
- Steiner, P.O. (1952), "Program patterns and the workability of competition in radio broadcasting." *Quarterly Journal of Economics*, 66(2), 194-223.
- Sweeting, A. (2010), "The Effects of Horizontal Mergers on Product Positioning: Evidence from the Music Radio Industry. *RAND Journal of Economics*, 41 (2): 372-97.
- Train, Kenneth (2009), *Discrete Choice Methods with Simulation*, Cambridge University Press, Second Edition.
- Weyl, E. G., "A price theory of multi-sided platforms." *American Economic Review*, 100(4): 1642-1672.
- Weyl, E. G. and A. White (2014), "Insulated Platform Competition." Working paper.
- Wilbur, K. C. (2008), "A Two-Sided, Empirical Model of Television Advertising and Viewing Markets," *Marketing Science*, 27 (3): 356-378.
- Zhou, Y. (2016), "Bayesian Estimation of a Dynamic Model of Two-Sided Markets: Application to the U.S. Video Games." *Management Science*, forthcoming.

Appendices

A.1 The Case $\kappa = 0$ Leads to a Binary Choice Model

In Section 4, I introduce a model of multiple discrete choices based on Hendel (1999) and Fan (2013). In this section, I show an interesting result that the multiple discrete choice model is equivalent to a binary choice model when $\kappa = 0$ and no restriction is imposed on the number of products that one can choose.

At the risk of abusing the notation, I denote \mathbb{P}_{ij}^m as the probability that product j is *exactly* consumer i 's m^{th} choice, and denote $\mathbb{P}_{ij}^{\geq m}$ as the probability that product j is *at least* consumer i 's m^{th} choice. Notice that I do not assume any specific functional form, so \mathbb{P}_{ij}^m and $\mathbb{P}_{ij}^{\geq m}$ are general probabilities. Observing that the events of j being one's first to m^{th} choice are mutually exclusive, I write $\mathbb{P}_{ij}^{\geq m}$ as

$$\mathbb{P}_{ij}^{\geq m} = \sum_{l=1}^m \mathbb{P}_{ij}^l. \quad (1)$$

Given that i has n choices from n products plus the no-purchase option, the probability i ever purchases j , \mathbb{P}_{ij} , is then

$$\begin{aligned} \mathbb{P}_{ij} &= \sum_{l=1}^n \mathbb{P}_{ij}^l = \sum_{l=1}^{n-1} \mathbb{P}_{ij}^l + \mathbb{P}_{ij}^n \\ &= \sum_{l=1}^{n-1} \mathbb{P}_{ij}^l + \underbrace{\left(\mathbb{P}_{ij}^{\geq n} - \sum_{l=1}^{n-1} \mathbb{P}_{ij}^l \right)}_{\text{by (24)}} \\ &= \mathbb{P}_{ij}^{\geq n}. \end{aligned}$$

In the context, $\mathbb{P}_{ij}^{\geq n}$ is the probability that i chooses j over the outside option. For example, with an i.i.d. Type I extreme value error and no random coefficients, $\mathbb{P}_{ij}^{\geq n}$ takes the binary logit form. This result is easy to understand intuitively. When a consumer faces n products and is free to choose up to n of them without diminishing utility after each choice, the consumer's choice problem reduces to saying yes-or-no

for each product. It further implies that each product is a monopoly to consumers, and specifically, all cross-price elasticities are automatically zero. If one is interested in studying substitution patterns and strategic firm behavior, this type of model may not be useful. In other contexts, binary choice models have been applied extensively. For example, Augereau, Greenstein, and Rysman (2008) uses a multivariate probit model to study R&D decisions; Hiedemann, Sovinsky and Stern (2013) uses a dynamic multivariate probit to study family long-term care decisions.

A.2 Market Penetration Function

In this section, I present details on the expression for the choice probability of household i ever choosing j . With the expression, it is easy to verify results on invertibility of the market penetration function, as shown in Fan (2013).

I first extend the notation used in Section 4. Let $\Phi_{ij}^{m,k}$ denote the probability j is i 's m^{th} best choice with utility decreased by k times; $\Phi_{ij-j^{(n)}}^{m,k}$ denotes the probability j is i 's m^{th} best choice when $j^{(n)}$ is not in the choice set and utility is decreased by k times. Here, $n = m - 1$. When $k = m - 1$, the superscript k is suppressed. For example, $\Phi_{ij}^1 = \Phi_{ij}^{1,0}$. Given that, I can write each of Φ_{ij}^n recursively as below:

$$\Phi_{ij}^2 = \sum_{j^{(1)} \neq j} \left(\Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1} \right); \quad (2)$$

where, as in the main text,

$$\Phi_{ij-j^{(1)}}^{1,1} = \frac{e^{\delta_j + \vartheta_{ij}}}{e^\kappa + \sum_{j \neq j^{(1)}} e^{\delta_h + \vartheta_{ih}}}, \quad \Phi_{ij}^{1,1} = \frac{e^{\delta_j + \vartheta_{ij}}}{e^\kappa + \sum_j e^{\delta_h + \vartheta_{ih}}}.$$

Similarly, I have

$$\Phi_{ij}^3 = \sum_{j^{(1)}, j^{(2)} \neq j} \left(\Phi_{ij-(j^{(1)}, j^{(2)})}^{1,2} - \Phi_{ij-j^{(1)}}^{2,2} - \Phi_{ij-j^{(2)}}^{2,2} - \Phi_{ij}^{1,2} \right); \quad (3)$$

$$\begin{aligned} \Phi_{ij}^4 = \sum_{j^{(1)}, j^{(2)}, j^{(3)} \neq j} & \left(\Phi_{ij-(j^{(1)}, j^{(2)}, j^{(3)})}^{1,3} - \Phi_{ij-(j^{(1)}, j^{(2)})}^{3,3} - \Phi_{ij-(j^{(1)}, j^{(3)})}^{3,3} - \Phi_{ij-(j^{(2)}, j^{(3)})}^{3,3} \right. \\ & \left. - \Phi_{ij-j^{(1)}}^{2,3} - \Phi_{ij-j^{(2)}}^{2,3} - \Phi_{ij-j^{(3)}}^{2,3} - \Phi_{ij}^{1,3} \right). \end{aligned}$$

Given parameters (κ, σ) , the market penetration function is

$$s_j(\delta, z, \nu; \kappa, \sigma) = \int \int \left(\Phi_{ij}^1(\delta, z, \nu; \sigma) + \sum_{n=2}^4 \Phi_{ij}^n(\delta, z, \nu; \kappa, \sigma) \right) dP_v dP_z. \quad (4)$$

Following BLP (1995), Fan (2013) shows that there exists a unique solution to $s_j(\delta, z, \nu; \kappa, \sigma) = \mathcal{S}_j$, where \mathcal{S}_j is data; and that $F_j = \delta_j + \ln \mathcal{S}_j - \ln s_j$ is a contraction mapping that can be used to invert s_j . In other words, this exists a one-to-one mapping between market penetration and mean product utilities. Essentially, it boils down to show: (C.1) $\partial s_j / \partial \delta_j < s_j$, (C.2) $\partial s_j / \partial \delta_j > 0$, (C.3) $\partial s_j / \partial \delta_h < 0$ for $h \neq j$, and (C.4) $\sum_{h=1, \dots, J} \partial s_j / \partial \delta_h > 0$. Together they imply the Jacobian of s has a dominate diagonal, therefore is a unique solution to $s_j(\delta, z, \nu; \kappa, \sigma) = \mathcal{S}_j$, and conditions in BLP (1995) met for F_j being a contraction mapping.

Intuitively, conditions (C.1) and (C.2) means that j 's penetration increases with its own quality, but the marginal effect cannot be “too large” (i.e., no marginal change in quality can double its sales). Condition (C.3) states products are substitutes. Condition (C.4) means j own quality effect dominates all cross-product quality effects. All of these carry economically sound intuitions, and can be easily verified with logistic functions.

First, I show that (C.1) holds. By definition, it is sufficient to show that $\frac{\partial \Phi_{ij}^n}{\partial \delta_j} <$

Φ_{ij}^n for each of the n . Due to the Type I Extreme Value errors, indeed, I have

$$\frac{\partial \Phi_{ij}^1}{\partial \delta_j} = \Phi_{ij}^1(1 - \Phi_{ij}^1) < \Phi_{ij}^1;$$

$$\begin{aligned} \frac{\partial \Phi_{ij}^2}{\partial \delta_j} &= \sum_{j^{(1)} \neq j} \left(\Phi_{ij-j^{(1)}}^{1,1}(1 - \Phi_{ij-j^{(1)}}^{1,1}) - \Phi_{ij}^{1,1}(1 - \Phi_{ij}^{1,1}) \right) \\ &< \sum_{j^{(1)} \neq j} (1 - \Phi_{ij-j^{(1)}}^{1,1}) \left(\Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1} \right) \\ &< \sum_{j^{(1)} \neq j} \left(\Phi_{ij-j^{(1)}}^{1,1} - \Phi_{ij}^{1,1} \right) = \Phi_{ij}^2; \end{aligned}$$

In addition, it can be further shown that $\frac{\partial \Phi_{ij}^2}{\partial \delta_j} < \Phi_{ij}^2(1 - \Phi_{ij}^2)$, which I shall use in the following steps.

$$\begin{aligned} \frac{\partial \Phi_{ij}^3}{\partial \delta_j} &< \sum_{j^{(1)}, j^{(2)} \neq j} \left(\Phi_{ij-(j^{(1)}, j^{(2)})}^{1,2}(1 - \Phi_{ij-(j^{(1)}, j^{(2)})}^{1,2}) - \Phi_{ij-j^{(1)}}^{2,2}(1 - \Phi_{ij-j^{(1)}}^{2,2}) \right. \\ &\quad \left. - \Phi_{ij-j^{(2)}}^{2,2}(1 - \Phi_{ij-j^{(2)}}^{2,2}) - \Phi_{ij}^{1,2}(1 - \Phi_{ij}^{1,2}) \right) \\ &< \sum_{j^{(1)}, j^{(2)} \neq j} (1 - \Phi_{ij-(j^{(1)}, j^{(2)})}^{1,2}) \left(\Phi_{ij-(j^{(1)}, j^{(2)})}^{1,2} - \Phi_{ij-j^{(1)}}^{2,2} - \Phi_{ij-j^{(2)}}^{2,2} - \Phi_{ij}^{1,2} \right) \\ &< \sum_{j^{(1)}, j^{(2)} \neq j} \left(\Phi_{ij-(j^{(1)}, j^{(2)})}^{1,2} - \Phi_{ij-j^{(1)}}^{2,2} - \Phi_{ij-j^{(2)}}^{2,2} - \Phi_{ij}^{1,2} \right) = \Phi_{ij}^3; \end{aligned}$$

$$\begin{aligned}
\frac{\partial \Phi_{ij}^4}{\partial \delta_j} &< \sum_{j^{(1)}, j^{(2)}, j^{(3)} \neq j} (\Phi_{ij-(j^{(1)}, j^{(2)}, j^{(3)})}^{1,3} (1 - \Phi_{ij-(j^{(1)}, j^{(2)}, j^{(3)})}^{1,3}) - \Phi_{ij-(j^{(1)}, j^{(2)})}^{3,3} (1 - \Phi_{ij-(j^{(1)}, j^{(2)})}^{3,3}) \\
&\quad - \Phi_{ij-(j^{(1)}, j^{(3)})}^{3,3} (1 - \Phi_{ij-(j^{(1)}, j^{(3)})}^{3,3}) - \Phi_{ij-(j^{(2)}, j^{(3)})}^{3,3} (1 - \Phi_{ij-(j^{(2)}, j^{(3)})}^{3,3}) \\
&\quad - \Phi_{ij-j^{(1)}}^{2,3} (1 - \Phi_{ij-j^{(1)}}^{2,3}) - \Phi_{ij-j^{(2)}}^{2,3} (1 - \Phi_{ij-j^{(2)}}^{2,3}) \\
&\quad - \Phi_{ij-j^{(3)}}^{2,3} (1 - \Phi_{ij-j^{(3)}}^{2,3}) - \Phi_{ij}^{1,3} (1 - \Phi_{ij}^{1,3})) \\
&< \sum_{j^{(1)}, j^{(2)}, j^{(3)} \neq j} (1 - \Phi_{ij-(j^{(1)}, j^{(2)}, j^{(3)})}^{1,3}) (\Phi_{ij-(j^{(1)}, j^{(2)}, j^{(3)})}^{1,3} - \Phi_{ij-(j^{(1)}, j^{(2)})}^{3,3} - \Phi_{ij-(j^{(1)}, j^{(3)})}^{3,3} \\
&\quad - \Phi_{ij-(j^{(2)}, j^{(3)})}^{3,3} - \Phi_{ij-j^{(1)}}^{2,3} - \Phi_{ij-j^{(2)}}^{2,3} - \Phi_{ij-j^{(3)}}^{2,3} - \Phi_{ij}^{1,3}) \\
&< \sum_{j^{(1)}, j^{(2)}, j^{(3)} \neq j} (\Phi_{ij-(j^{(1)}, j^{(2)}, j^{(3)})}^{1,3} - \Phi_{ij-(j^{(1)}, j^{(2)})}^{3,3} - \Phi_{ij-(j^{(1)}, j^{(3)})}^{3,3} \\
&\quad - \Phi_{ij-(j^{(2)}, j^{(3)})}^{3,3} - \Phi_{ij-j^{(1)}}^{2,3} - \Phi_{ij-j^{(2)}}^{2,3} - \Phi_{ij-j^{(3)}}^{2,3} - \Phi_{ij}^{1,3}) = \Phi_{ij}^4.
\end{aligned}$$

Therefore, it follows that

$$\begin{aligned}
\frac{\partial s_j}{\partial \delta_j} &= \int \int \sum_{n=1}^4 \left(\frac{\partial \Phi_{ij}^n}{\partial \delta_j} \right) dP_v dP_z \\
&< \int \int \sum_{n=1}^4 (\Phi_{ij}^n) dP_v dP_z = s_j.
\end{aligned}$$

I now turn to show that (C.3) holds; that is, the market share of j is decreasing in the mean product utility of h .

To begin with, observe that $\frac{\partial \Phi_{ij}^1}{\partial \delta_h} = -\Phi_{ij}^1 \Phi_{ih}^1 < 0$. Using that, I find

$$\begin{aligned}
\frac{\partial \Phi_{ij}^1}{\partial \delta_h} + \frac{\partial \Phi_{ij}^2}{\partial \delta_h} &= -\Phi_{ij}^1 \Phi_{ih}^1 + \Phi_{ij}^{1,1} \Phi_{ih}^{1,1} + \sum_{j^{(1)} \neq j, h} \left(-\Phi_{ij-j^{(1)}}^{1,1} \Phi_{ih-j^{(1)}}^{1,1} + \Phi_{ij}^{1,1} \Phi_{ih}^{1,1} \right) \\
&< \sum_{j^{(1)} \neq j, h} \left(-\Phi_{ij-j^{(1)}}^{1,1} \Phi_{ih-j^{(1)}}^{1,1} + \Phi_{ij}^{1,1} \Phi_{ih}^{1,1} \right) < 0.
\end{aligned}$$

Recall that, in my notation, $\Phi_{ij}^1 = \Phi_{ij}^{1,0}$ and $\Phi_{ij}^2 = \Phi_{ij}^{2,1}$, so the above inequality does actually hold for any pair of $k-1$ and k . One can then match the corresponding pairs to show that $\frac{\partial \Phi_{ij}^1}{\partial \delta_h} + \frac{\partial \Phi_{ij}^2}{\partial \delta_h} + \frac{\partial \Phi_{ij}^3}{\partial \delta_h} < 0$, and $\frac{\partial \Phi_{ij}^1}{\partial \delta_h} + \frac{\partial \Phi_{ij}^2}{\partial \delta_h} + \frac{\partial \Phi_{ij}^3}{\partial \delta_h} + \frac{\partial \Phi_{ij}^4}{\partial \delta_h}$, inductively.

It then follows that $\frac{\partial s_j}{\partial \delta_j} = \int \int \sum_{n=1}^4 \left(\frac{\partial \Phi_{ij}^n}{\partial \delta_h} \right) dP_v dP_z < 0$.

Next is (C.4). As in Fan (2013), note that $\sum_{h=1 \dots J} \partial s_j / \partial \delta_h = \frac{\partial s_j(\delta + \Delta)}{\partial \Delta} \Big|_{\Delta=0}$.

Since

$$\begin{aligned} \frac{\partial \Phi_{ij}^1(\delta + \Delta)}{\partial \Delta} \Big|_{\Delta=0} &= (\Phi_{ij}^1)^2 \frac{1}{e^{\delta_j + \vartheta_{ij}}} > 0; \\ \frac{\partial \Phi_{ij}^2(\delta + \Delta)}{\partial \Delta} \Big|_{\Delta=0} &= \sum_{j^{(1)} \neq j} \left((\Phi_{ij-j^{(1)}}^{1,1})^2 - (\Phi_{ij}^{1,1})^2 \right) \frac{e^\kappa}{e^{\delta_j + \vartheta_{ij}}} > 0; \end{aligned}$$

$$\begin{aligned} \frac{\partial \Phi_{ij}^3(\delta + \Delta)}{\partial \Delta} \Big|_{\Delta=0} &= \sum_{j^{(1)}, j^{(2)} \neq j} \left((\Phi_{ij-(j^{(1)}, j^{(2)})}^{1,2})^2 - (\Phi_{ij-j^{(1)}}^{2,2})^2 \right. \\ &\quad \left. - (\Phi_{ij-j^{(2)}}^{2,2})^2 - (\Phi_{ij}^{1,2})^2 \right) \frac{e^{2\kappa}}{e^{\delta_j + \vartheta_{ij}}} > 0; \end{aligned}$$

and,

$$\begin{aligned} \frac{\partial \Phi_{ij}^4(\delta + \Delta)}{\partial \Delta} \Big|_{\Delta=0} &= \sum_{j^{(1)}, j^{(2)}, j^{(3)} \neq j} \left((\Phi_{ij-(j^{(1)}, j^{(2)}, j^{(3)})}^{1,3})^2 - (\Phi_{ij-(j^{(1)}, j^{(2)})}^{3,3})^2 - (\Phi_{ij-(j^{(1)}, j^{(3)})}^{3,3})^2 \right. \\ &\quad \left. - (\Phi_{ij-(j^{(2)}, j^{(3)})}^{3,3})^2 - (\Phi_{ij-j^{(1)}}^{2,3})^2 - (\Phi_{ij-j^{(2)}}^{2,3})^2 \right. \\ &\quad \left. - (\Phi_{ij-j^{(3)}}^{2,3})^2 - (\Phi_{ij}^{1,3})^2 \right) \frac{e^{3\kappa}}{e^{\delta_j + \vartheta_{ij}}} > 0. \end{aligned}$$

It follows immediately that

$$\sum_{h=1 \dots J} \partial s_j / \partial \delta_h = \frac{\partial s_j(\delta + \Delta)}{\partial \Delta} \Big|_{\Delta=0} = \int \int \sum_{n=1}^4 \left(\frac{\partial \Phi_{ij}^n(\delta + \Delta)}{\partial \Delta} \Big|_{\Delta=0} \right) dP_v dP_z > 0.$$

Together, (C.3) and (C.4) imply (C.2). Thus, all four conditions are verified.

A.3 Data

A.3.1 Sample of Magazines

My data set on U.S. magazines comes from a number of sources, described in the next subsection. I merge various data to create a panel of magazines. I delete magazines with important variables missing. I also delete any genre of magazine in which at least one important magazine is missing or deleted from the data. The result is a sample of 34 major magazines in six genres for the main analysis. For various robustness analyzes, I include 30 more magazines that have complete information on advertising price and quantity.

Table 1: The List of Magazines for Main Analysis

Women's fashion	Women's health
<i>Allure</i>	<i>Fitness</i>
<i>Cosmopolitan</i>	<i>Health</i>
<i>Elle</i>	<i>Prevention</i>
<i>Glamour</i>	<i>Self</i>
<i>Harper's Bazaar</i>	
<i>In Style</i>	Shelter
<i>Lucky</i>	<i>Architectural Digest</i>
<i>Marie Claire</i>	<i>Country Living</i>
<i>More</i>	<i>This Old House</i>
<i>Seventeen</i>	<i>Town & Country</i>
<i>Vogue</i>	
<i>W</i>	Man's
	<i>Details</i>
Women's general	<i>Esquire</i>
<i>Family Circle</i>	<i>Maxim</i>
<i>Good Housekeeping</i>	<i>Playboy</i>
<i>Martha Stewart Living</i>	
<i>Real Simple</i>	Business
<i>Redbook</i>	<i>Money</i>
<i>Woman's Day</i>	<i>Fast Company</i>
	<i>Forbes</i>
	<i>Entrepreneur</i>

A.3.2 Data Sources

Below, Table 2 summarizes data sources and definition.

Table 2: Data Description and Sources

	Var	Data description	Data source
Magazine demand	N_{jct}^s	MSA-level circulation	AAM
Ad demand	a_{jt}	Annual number of advertising pages	PIB-Kantar
Price of magazine	p_{jt}^s	Annual subscription price (2012 \$)	AAM
Price of ad per page	p_{jt}^a	Average advertising rate (2012 \$/page)	PIB-Kantar
Magazine characteristics	x_{1jt}	Annual number of non-ad content pages	MA-Focus
	x_{2jt}	Frequency of publication (issues/year)	AAM
Consumer rankings	Φ_{jt}^n	% consumer ranking groups	GfK MRI
Ownership		Publisher	AAM
MSA demographics	\mathcal{M}_{jct}	Number of households	AAM
	z_{1ict}	Log HH income	Census
	z_{2ict}	Age	Census
	z_{3ict}	Household sex ratio	
	z_{4ict}	Education (years)	Census
	z_{5ict}	Homeownership (=1 if owned, =0 otherwise)	Census
	z_{6ict}	Migration (=1 if recently moved)	Census

Notes: AAM: Alliance of Audited Media, formerly Audit Bureau of Circulation (ABC); PIB-Kantar: Publishing Information Bureau; MA-Focus: MA-focus media, formerly Hall's report; GfK MRI: ; Census: American Community Survey.

A.3.3 Cross-Sectional Variation of Magazine Penetration

Table 3 presents the summary statistics of market penetration of each magazine in 2012. It shows that market penetration of the same magazine varies substantially across metro areas in 2012. Similar patterns (not reported) are present in data of other years.

Table 3: Summary Statistics of Magazine Penetration in 2012

	Mean	SD	Min	Max
Magazine 1	0.77	0.21	0.38	1.68
Magazine 2	0.53	0.23	0.21	2.52
Magazine 3	2.45	0.08	0.83	8.82
Magazine 4	1.56	0.50	0.26	3.24
Magazine 5	0.81	0.38	0.31	6.51
Magazine 6	0.52	0.13	0.16	1.21
Magazine 7	0.47	0.15	0.21	1.15
Magazine 8	3.68	0.90	1.34	6.54
Magazine 9	0.59	0.13	0.32	1.58
Magazine 10	1.20	0.28	0.73	2.92
Magazine 11	0.74	0.16	0.38	1.72
Magazine 12	1.83	0.38	1.13	4.07
Magazine 13	3.93	0.94	1.26	6.74
Magazine 14	0.45	0.19	0.23	1.77
Magazine 15	1.17	0.25	0.52	2.26
Magazine 16	1.22	0.52	0.36	4.70
Magazine 17	0.80	0.22	0.42	1.81
Magazine 18	0.69	0.16	0.39	1.38
Magazine 19	1.67	0.43	0.56	4.13
Magazine 20	2.11	0.44	0.54	4.43
Magazine 21	1.50	0.45	0.55	3.26
Magazine 22	1.02	0.28	0.34	2.13
Magazine 23	2.51	0.44	1.18	4.09
Magazine 24	1.56	0.56	0.39	4.07
Magazine 25	1.93	0.48	0.85	3.41
Magazine 26	1.16	0.36	0.50	5.58
Magazine 27	1.58	0.42	0.81	3.11
Magazine 28	0.83	0.31	0.25	1.96
Magazine 29	0.31	0.15	0.11	1.45
Magazine 30	0.85	0.39	0.35	4.11
Magazine 31	0.30	0.11	0.16	0.87
Magazine 32	3.66	0.97	0.78	8.75
Magazine 33	0.31	0.09	0.15	0.77
Magazine 34	3.04	0.68	1.04	5.25