Essays on Industrial Organization

(Article begins on next page)

Dissertation Advisor: Ariel Pakes   Michael James Sinkinson

Essays on Industrial Organization

ABSTRACT

This dissertation consists of three essays in the area of Industrial Organization.

The first essay established the theoretical motivations for, and implications of, exclusive contracts, with an application to smartphones. Why would Apple choose to distribute its smartphone through only one carrier, and why would AT&T bid the most for exclusivity? I develop a model which shows that if upstream handset manufacturers face a relatively low price elasticity for their good compared to downstream wireless carriers, exclusive contracts can maximize their joint profits. An exclusive contract reduces price competition in the final good market but also increases returns to innovation for parties outside the contract, such as Google's Android. Different price elasticities among downstream firms due to network quality differences lead to different valuations of the exclusive contract.

The second essay estimates the relative elasticities of smartphone and carrier demand using simulation and MCMC methods on a detailed monthly dataset of consumer decisions over 2008-2010. Counterfactual simulations show the importance of recomputing the price equilibrium to understanding the observed market structure. Accounting for price effects, AT&T had the

highest value of exclusivity with Apple, and was willing to compensate Apple $148 per unit sale foregone. Apple's exclusivity increased entry incentives for Android handset makers by approximately $1B.

The third essay uses data on US newspapers from the early 20th century to study the economic incentives that shape ideological diversity in the media. My co-authors and I show that households prefer newspapers whose political content agrees with their own ideology, that newspapers with the same political content are closer substitutes than newspapers with different political content, and that newspapers seek both to cater to household tastes and to differentiate from their competitors. We estimate a model of newspaper demand, entry, and affiliation choice that captures these forces. We show that competitive incentives greatly enhance the extent of ideological diversity in local news markets, and we evaluate the impact of policies designed to increase such diversity.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

1. INTRODUCTION

This dissertation consists of three papers relating to the field of Industrial Organization. The first two examine the mobile telecommunications industry, first through a theoretical examination of vertical contracting, and then through an empirical study of the forces that have shaped the market. The third looks at the setting of newspapers, performing an empirical investigation into the competing forces that shape the observed diversity of news content available. The common thread throughout this work is the focus on classic questions of the Industrial Organization literature, such as pricing, contracting, entry, and product positioning, as they apply to industries of great policy relevance.

The first paper establishes a theoretical motivation for exclusive contracts between upstream firms and downstream firms when a final good consists of both an upstream and downstream input, and there is horizontal differentiation at both levels. One example of such a setting is wireless telecommunications, where upstream goods, handsets, are bundled with downstream goods, wireless networks, to form a final good. In such a setting, an exclusive contract can maximize the joint profits of the contracting parties when handsets are poor substitutes for one another, but wireless networks are good substitutes for one another. The reason for this is that when prices of goods are strategic complements, as is typical in horizontally differentiated markets, the exclusive contract moves the price in the final goods market to a higher price equilibrium. I provide the conditions under which the exclusive contract

is optimal for the contracting parties, and further explore the implications for such a contract to parties outside the contract. Finally I investigate how the willingness to pay for such a contract will differ by the demand elasticity faced by the downstream firms.

The second paper estimates demand for smartphones and wireless networks using a rich dataset of consumer choices with the goal of estimating the strength of the mechanism explored in the first paper. The demand dataset consists of 26 months of a 25,000-person repeated cross sectional survey of the United States wireless market. I augment this with detailed measures of network quality for all of the major carriers in 90 markets, as well as additional product characteristic data. I develop a model of consumer demand and estimate the model using simulation methods. Finally, I perform counterfactual simulations to determine the incentives present for Apple to release its iPhone exclusively on a single carrier and to quantify the implications of that contract for other parties.

The third paper, which is joint work with Matthew Gentzkow and Jesse M. Shapiro of the University of Chicago Booth School of Business, measures the competing forces that determine product diversity in the newspaper industry. The classical tension in the product positioning literature is between locating near demand versus far from my competition. We measure these forces in terms of a newspaper's political affiliation using a historical dataset of all US daily newspapers. Our method exploits spatial correlation in unobservable preferences to separately identify the two forces. We find that the need to differentiate from competitors is an important factor in determining the diversity of news viewpoints available in markets. Through counterfactual simulations, we are able to analyze the impact of different policy interventions meant to increase the diversity of news viewpoints available.

2. THEORETICAL MOTIVATION FOR EXCLUSIVE CONTRACTS IN MULTIBUYER-MULTISELLER SETTINGS

2.1 Introduction

A final good in the smartphone market consists of both a smartphone handset and the wireless service that enables it to function. Exclusive contracts in this market between upstream firms (handset manufacturers) and downstream firms (wireless carriers) are common.[1] Perhaps the most well-known is the contract between Apple and AT&T, which saw the former's iPhone handset exclusively available on AT&T's network in the United States. An exclusive contract such as this restricts Apple from engaging in trade with competing wireless carriers, and so the contract must compensate Apple for the lost market potential. Early models of exclusive contracts argued that such arrangements must be efficient, as AT&T would only be willing to sufficiently compensate Apple for the lost sales if the exclusive arrangement was efficient.[2] However, later approaches showed that such arrangements could lead to inefficient outcomes, such as the foreclosure of entry (Aghion and Bolton, 1987). While these contracts may have anti-competitive effects, they have also been shown to be

---

[1] For example, in Consumer Reports' 2009 annual review of smartphones, 6 of the 10 devices that were rated as "Recommended" were exclusive to one of the four major US wireless carriers. See Consumers Union of US (2009): "CELL PHONES: Our tests of 70 standard and smart models show theyre sharing many more features," Consumer Reports, 74(1), Albany, NY.

[2] These arguments, referred to as the Chicago School approach to this topic, are articulated in Posner (1976) and Bork (1978).

pro-competitive in some settings, such as for protecting investments and addressing externalities (Bernheim and Whinston, 1998; Segal and Whinston, 2000).[3] Indeed, courts in the United States evaluate non-price vertical restraints under the Rule of Reason, instead of declaring them to be illegal *per se*.[4]

This paper proposes a simple motivation for exclusivity in the mobile telecommunications market based on the relative substitutability of the upstream goods (handsets) versus the downstream goods (wireless service). If the downstream goods are near-perfect substitutes, then downstream firms face high price elasticities for their goods and are only able to charge low markups above marginal cost for their goods in equilibrium. If the upstream goods are poor substitutes, those firms face low price elasticities and are able to charge large markups over marginal cost in equilibrium. I show that in such a setting, an exclusive contract can maximize the joint profits of the contracting parties by reducing price competition in the final goods market. However, these contracts also increase incentives for new upstream firms to enter. Finally, I investigate the willingness to pay of differentiated downstream firms, and find that firms with lower quality goods may be willing to bid the most for an exclusive contract as they have more to lose from a rival gaining exclusivity.

Some alternative mechanisms have been put forward to explain Apple's choice to enter into an exclusive contract. A first such argument was that Apple had a limited supply capacity: this was their first mobile phone, and so they were concerned that they could not meet demand if they launched on all carriers. However, if this were the case, it is unlikely that they would then have entered into a 5-year exclusive contract. Apple launched the iPhone globally less than 6 months after the initial US launch,

---

[3] See Katz (1989) for a survey of the literature on vertical contracts.

[4] United States Supreme Court, "433 U.S. 36 CONTINENTAL T. V., INC., ET AL. v. GTE SYLVANIA INC."

indicating that any supply issues were short-term. A second argument was that exclusivity was essential to guarantee carrier investments in network technologies to support the iPhone. However, this argument was specifically rejected by the French competition authorities when they prematurely ended Apple's exclusive contract in that country. The exclusive carrier there was unable to show a significant investment that needed to be protected.[5]

This paper consists of a theoretical analysis of firm decisions in this setting. The theory model builds on the approach taken by Rey and Stiglitz (1995), which shows that upstream competition can lead to exclusive contracts with undifferentiated downstream firms if the upstream goods are imperfect substitutes and prices are strategic complements.[6] The mechanism is that exclusivity decreases the interbrand price competition among upstream firms. This is more closely related to the setting at hand, as handsets are horizontally differentiated. However, the result also relies on the downstream firms being perfect substitutes and having no market power. This paper contributes a more general model that allows for downstream horizontal differentiation. I find that when upstream demand is relatively less sensitive to price than downstream demand, exclusive contracts can lessen price competition and overcome the losses associated with being available with fewer downstream firms. Furthermore, if downstream firms face different price elasticities for their goods, their willingness to pay for exclusivity will differ. I show that if consumers are willing to substitute between handset and network quality, a lower quality carrier may benefit more from an exclusive contract. Finally, I show that the existence of exclusive contracts can

---

[5] Conseil de la concurrence: Décision n. 08-MC-01 du 17 décembre 2008 relative à des pratiques mises en oeuvre dans la distribution des iPhones.

[6] If the prices of two firms' products are strategic complements, then an increase in the price of one good gives the other firm an incentive to increase the price of the other good as well. See Bulow, Geanakoplos, and Klemperer (1985).

increase entry incentives for parties outside of the contract.

This paper's contributions to the literature are an extension of the theoretical understanding of exclusive contracting to the case of horizontal differentiation at both the upstream and downstream levels and an investigation of the implications of exclusive contracts in such a setting.

The paper proceeds as follows: Section 2 discusses the setting in more depth. Section 3 provides an illustrative example. Section 4 generalizes the model. Section 5 summarizes. All proofs are found in the Appendix.

2.2 *Model Overview*

The setting in question is one where upstream firms (say, handset manufacturers) sell a good to downstream firms (say, wireless carriers), who bundle this good with their own product and sell the final bundle to consumers. While models of vertical settings are common in economic theory, most models are limited to "triangular" market structures, with either one upstream firm and two downstream firms, or vice versa.[7] This section begins with an example where downstream goods are homogeneous to illustrate the static incentive for exclusive contracts in a simplified setting. Specifically, exclusive contracts lead to steeper reaction functions for the upstream firms, resulting in higher prices in equilibrium. The model is then generalized to allow for differentiated goods at both levels, to match the reality of the US mobile telecommunications industry and establish the main theoretical results. The main findings are that exclusivity is optimal when the downstream goods are good sub-

---

[7] Whinston (2006) notes this and further states that most markets in reality have multiple participants at each level. One exception is Besanko and Perry (1994), which has two upstream firms and multiple downstream firms spatially differentiated as in a Salop circle model. However, the contracts are restricted to be linear and an exclusive contract in their setting only restricts the upstream competitor from every 2nd downstream firm.

stitutes for one another, that exclusive contracts can lead to entry that would not be profitable in their absence, and that the value of the exclusive contract to a downstream firm depends on whether consumers are willing to substitute between quality of the upstream and downstream goods.

The specific terms of vertical contracts are unobserved in the mobile telecommunications sector, and so I wish to abstract away from bargaining over surplus between the contract parties. Instead I look at the joint surplus of the contracting parties as the determinant of the market structure. This is consistent with other research on exclusivity, such as Bernheim and Whinston (1998).[8,9] I will refer to the case of non-exclusivity as *common agency*, denoted by *C* below, the case of single-firm exclusivity as *E*, and of all upstream firms exclusive by *EE*.

2.3   Illustrative Example

An important distinction in this setting is the fact that a new smartphone is an imperfect substitute for an existing one; that is, while a given consumer may prefer an iPhone to, say, a Blackberry, there exists a set of prices at which the consumer would prefer the Blackberry. This imperfect competition allows for a static motivation for exclusive contracts.

Consider a simplified static setup (see Appendix A.1 for all derivations): Firm A could invest K to develop a new smartphone. If it enters the market, it would have a smartphone with quality δ_A and marginal cost c, that would compete against Firm B

---

[8] The first principle from Bernheim & Whinston's analysis of manufacturers and exclusive retailers: "the form of representation (exclusivity or common representation) that arises in equilibrium maximizes the joint surplus of the manufacturers and the retailer, subject to whatever inefficiencies may (or may not) characterize incentive contracting between the retailer and the manufacturers."

[9] To this end, I will allow for flexible contracts so that classic results such as double-marginalization are not an issue.

that produces a smartphone with quality δ_B at marginal cost c. Consumer tastes for smartphones are as in a standard Hotelling model where consumers are distributed uniformly over an interval of length 1, with tastes for each smartphone for consumer i at location θ_i given by:

$$
\begin{aligned}
u_{Ai} &= \delta_A - p_A - \theta_i \\
u_{Bi} &= \delta_B - p_B - (1 - \theta_i)
\end{aligned}
$$

The smartphones are purchased from the manufacturers at wholesale prices q_A and q_B by N identical wireless carriers. These carriers compete in the downstream market by bundling the devices with their homogeneous wireless networks that have marginal cost of zero, and selling the handset-network bundle to consumers at prices p_A and p_B. See Figure 2.1 for a diagram of this setup. Appendix A.1, shows the derivation of final consumer demand as a function of prices, $D^A(p_A, p_B)$ and $D^B(p_A, p_B)$, by locating the indifferent consumer and using the properties of the uniform distribution, as is standard for a Hotelling setup.

Firm A could choose to sell its handset to all carriers, or limit itself to a single exclusive carrier. I will first hold Firm B's choice fixed at non-exclusivity for now, but will revisit Firm B's choice at the end. I begin by analyzing Firm A's expected profits from common-agency, followed by the profits from exclusivity. The order of moves for this full-information setup is (1) upstream firms simultaneously choose wholesale prices, (2) carriers simultaneously choose retail prices, and (3) the market is realized.[10]

If no exclusive contracts are permitted, then all carriers will offer a bundle with

---

[10] Given the full-information setup of the game, the sequential nature merely facilitates exposition.
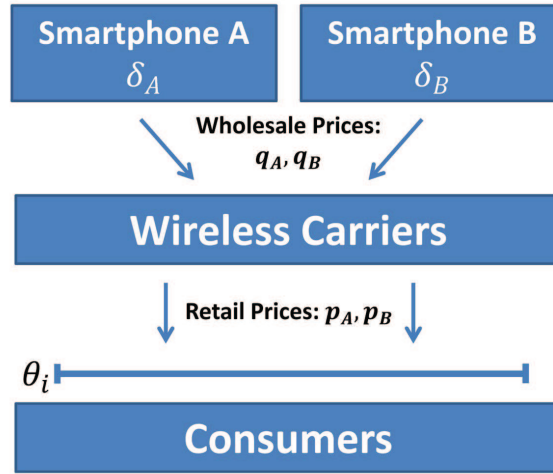
Figure 2.1: Illustrative Model Setup

each smartphone, and Bertrand competition will ensure that markups are competed to zero. Knowing this, the smartphone firms will choose wholesale prices in equilibrium to maximize their profits given that the downstream firms will not charge a markup:

$$\pi_A^c = (q_A - c) D^A (q_A, q_B)$$
$$\pi_B^c = (q_B - c) D^B (q_A, q_B)$$

Assuming an interior solution,[11] the equilibrium wholesale price and profits for firm A if it enters with no exclusive arrangement are π_A^{C*}, shown in Table 2.1 with the resulting retail price. This is identical to the level profits earned if the two smartphone firms competed directly for consumers, due to Bertrand competition among the homogeneous carriers.

---

[11] Interior refers to the case where δ^A and δ^B are such that neither firm captures the entire market in equilibrium.

Now suppose that Firm A could instead sign an agreement with one carrier guaranteeing exclusivity: Firm A could not sell its smartphone to any other carrier, but the carrier would be free to offer smartphone B. This is more closely aligned with the concept of "exclusive territories" than "exclusive contracts" in the literature (Katz, 1989). In this case, Firm A would expect its exclusive wireless carrier w to choose a retail price to maximize profits, where the carrier's profits and optimal retail price are given by:

$$
\begin{aligned}
\pi_w^E &= (p_A - q_A) D^A (p_A, q_B) \\
p_A^{E*} &= \left(\frac{1 + \delta_A - \delta_B + p_B + q_A}{2} \right)
\end{aligned}
$$

The upstream firms choose wholesale prices knowing this markup. Upstream profits[12] are now

$$
\begin{aligned}
\pi_A^E &= \left(p_A^{E*} (q_A, q_B) - c \right) D^A \left(p_A^{E*} (q_A, q_B), q_B \right) \\
\pi_B^E &= (q_B - c) D^B \left(p_A^{E*} (q_A, q_B), q_B \right)
\end{aligned}
$$

Solving for equilibrium wholesale prices, we see that Firm B reaction function now takes the downstream optimization into account, and so is more inelastic with respect to Firm A's wholesale price. Consequently, both smartphones have higher prices over the range of interior solutions. Firm A's profit under exclusivity π_A^{E*}, is greater than its profits under common agency.

If Firm B were also exclusive, both firms would internalize the downstream pricing behavior, and Firm A's profits from exclusivity would rise further. Table 2.1

[12] Note that Firm A's profits include the downstream firm's markup. It is assumed that when exclusive, upstream firms are able to extract the full surplus via a fixed fee in a two-part tariff.

Form of Representation	Retail Price, A	Profits, Firm A
Common Agency (C)	$c + 1 + \frac{1}{3}(\delta_A - \delta_B)$	$\frac{1}{18}(3 + \delta_A - \delta_B)^2$
A Exclusive (E)	$c + \frac{5}{4} + \frac{1}{4}(\delta_A - \delta_B)$	$\frac{1}{32}(5 + \delta_A - \delta_B)^2$
A, B Exclusive (EE)	$c + 2 + \frac{2}{5}(\delta_A - \delta_B)$	$\frac{1}{25}(5 + \delta_A - \delta_B)^2$

summarizes the outcomes of this setup.

We may now draw a few conclusions from this model:

1. Firm A will earn greater profits under exclusivity. This result is not particularly novel: Rey and Stiglitz (1995) proved this in the setting of producers and retailers for a general quasi-concave profit function where $\delta_A = \delta_B$ and both upstream firms move simultaneously. Their Proposition 3 states that if retail prices are strategic complements and profit functions are quasi-concave, then both smartphone firms would choose exclusivity. The model described above meets their criteria.

2. There exist values of K such that a rational Firm A would choose not to enter in the absence of exclusive contracts. Furthermore, if the incumbent is exclusive, the entry incentive is even greater when exclusive contracts are available. This is a direct result of the above, but is interesting in that it provides evidence that exclusive contracts increase the returns to innovation.

What is driving this result? A major force at work is that downstream Bertrand competition drives markups to zero, and so exclusivity provides a buffer against price competition. The exclusive contract alters the response curves of the upstream firms, taking advantage of the fact that prices are strategic complements. Below I will extend the general model to the case of differentiated goods at both upstream and downstream levels and show that under certain conditions, exclusivity is the

optimal contract. In many realistic settings, downstream firms are differentiated or contributed a differentiated good to the end product, and so this generalization is relevant.[13]

2.4 *General Model*

We can think of the case above as a limit case where downstream firms are perfect substitutes to consumers. Another limit case is where downstream firms are not substitutes at all, or where wireless carriers are effectively monopolists over their customers. In that setting, it is clear that exclusivity can not be optimal for an upstream firm, as they could do strictly better selling to 2 or more downstream firms, as each carrier is effectively a separate market. Figure 2.2 illustrates the profits to the entering upstream firm at different levels of downstream market power, and for different contracts, providing a roadmap to this section. I maintain throughout the assumption that competing handsets are imperfect substitutes and that prices of handsets are strategic complements. For simplicity, I will assume that the underlying demand system captures downstream "substitutability" with a parameter $\eta \in [0, \infty)$, such that under common agency, when $\eta = 0$, downstream firms are perfect substitutes as in the above section so that for carrier n , where s_{An} is the share of handset A on carrier n, we have that $\frac{\partial s_{An}}{\partial p_{An}} = -\infty$. As η increases, so does $\frac{\partial s_{An}}{\partial p_{An}}$, and in the limit $\frac{\partial s_{An}}{\partial p_{An}} \to \frac{\partial s_A}{\partial p_A}$ as $\eta \to \infty$. This allows us to characterize the limit cases of carrier monopolists ($\eta = \infty$), carriers as perfect substitutes ($\eta = 0$), and cases in-between. As an example of how such a parameterization could arise, consider a standard Hotelling setup where the transport cost across the unit interval is given by η: when $\eta = 0$, all

---

[13] Whinston (2006) states with regard to multibuyer/multiseller settings that "developing models that reflect this reality is a high priority."
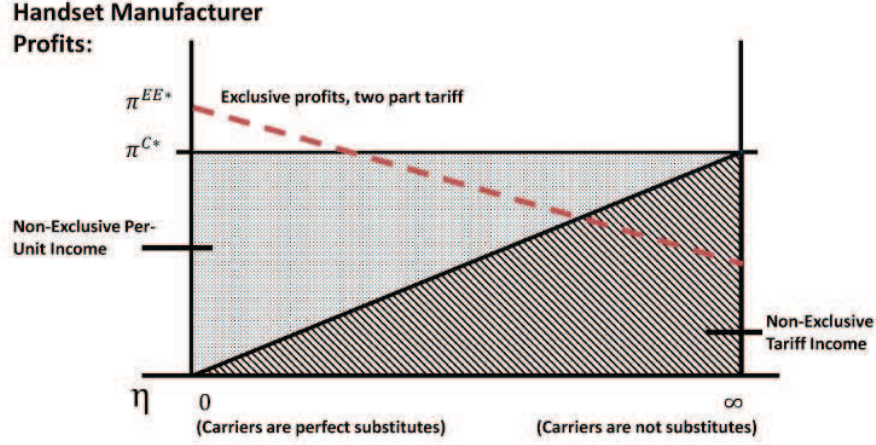
Figure 2.2: Upstream Firm Profits by Contract and Downstream Market Power

consumers are equally willing to go to either end of the interval, and as η increases, consumers are less willing to substitute to the firm that is located further from them. Appendix A.2 details additional examples of demand systems with this property.

We will now consider the general case of two upstream firms as before, but now N downstream firms that are imperfect substitutes. Under non-exclusivity for both A and B, the maximum possible profits for firm A under a two-part tariff are given by the profits earned from selling directly to consumers:

$$\pi_A^{C*} = \frac{s_A \left(p_A^{C*}, p_B^{C*} \right)^2}{-\frac{\partial s_A}{\partial p_A}}$$

The details of how this is achieved at any η are in Appendix A.2.

Under exclusivity, carriers 1 and 2 have exclusivity of products A and B respectively, and choose markups based on the wholesale prices they are charged. It is easy to show that these markups are greater than the markups they choose under com-

mon agency at a given wholesale price. Knowing the expected markup functions, the handset makers choose wholesale prices to maximize their joint profits with their exclusive carrier. This yields a best response function for each of the handset makers that is far steeper than the common-agency setting. Let $m_h(q_A, q_B)$ denote the carrier's markup function for handset h, and note that it is decreasing in own wholesale price but increasing in opposite wholesale price. We have a best response function for Firm A of

$$q_A - c = -m_A + \frac{\left(1 + \frac{\partial m_A}{\partial q_A}\right) s_{A1}}{-\left(\frac{\partial s_{A1}}{\partial p_{A1}}\left(1 + \frac{\partial m_A}{\partial q_A}\right) + \frac{\partial s_{A1}}{\partial p_{B2}}\frac{\partial m_B}{\partial q_A}\right)}$$

We see that the handset maker effectively replaces the carrier's markup with a more optimal one, which is based on a lower elasticity when prices are strategic complements (as captured by $\frac{\partial s_{A1}}{\partial p_{B2}}\frac{\partial m_B}{\partial q_B}$). This results in a higher retail price for both handsets, and profits under exclusivity of π_A^{EE*}. Figure 2.2 summarizes the upstream profits under different contract forms at different levels of downstream market power.

We can now turn to our first result:

Proposition 1. *In the above model, if (a) prices are strategic complements, (b) shares are smooth and twice continuously differentiable in prices, (c) the price equilibrium exists, is unique, and continuous, then there exists a value η^* such that for all $\eta < \eta^*$, exclusivity is jointly profit maximizing.*

The proof follows from the fact that final retail prices are higher under exclusivity, but market share is lower (except in the case of carriers as perfect substitutes). The formal proof relies on continuity and the Intermediate Value Theorem, since $\pi_A^{EE*}(\eta = 0) > \pi_A^{C*}$, but $\pi_A^{EE*}(\eta = \infty) < \pi_A^{C*}$. From the proof, we can see that the range of downstream elasticity over which exclusivity is optimal is (a) decreasing with N, the number of wireless carriers, (b) increasing with the degree of comple-

14

mentarity of prices, and (c) decreasing with the elasticity of upstream demand. These are all intuitive findings: the first captures the fact that as the number of downstream firms increases, so does the opportunity cost of exclusivity. The second captures the degree of the pricing advantage of exclusive contracting, and the third captures the influence exclusivity will have on downstream market shares.

Corollary. *The existence of exclusive contracts can lead to entry in cases where it would not be profitable otherwise.*

This is a direct consequence of the above proposition. There is a non-empty range of entry costs such that entry is not profitable in the absence of exclusive contracts, but is profitable with exclusivity.

Until now we have considered downstream firms to be identical and horizontally differentiated. Suppose now that for simplicity there are only two downstream firms (carriers) and that they also differ in a vertical characteristic. One example of this for wireless carriers could be the quality of their network (e.g. dropped call rate). Suppose further that a handset maker has decided to enter exclusively. When might we expect one carrier or the other to be the most profitable match for exclusivity? Assume that a carrier would be willing to pay up to its profit difference between exclusivity and rival exclusivity (i.e. AT&T would have been willing to pay Apple up to its profit difference between AT&T-Apple exclusivity and Verizon-Apple exclusivity).

Based on the model above, it seems intuitive that a carrier that faces more elastic demand would have the most to lose from a rival gaining exclusivity, as it would face a larger change in equilibrium price. Assume that consumers observe a vertical characteristic of each carrier n, δ_n, with $\delta_n \neq \delta_{n'}$ and price elasticity at a given price decreasing in δ_n. Further assume that consumer utility for the handset-network bundle (δ_A, δ_n) takes the form $u_{An} = \delta_A + \delta_n + \beta \delta_A \delta_n - p_{An}$. This form is chosen as

the interaction term allows consumers to "substitute" between handset and network quality ($\beta < 0$), or it allows a better network to make a handset even better ($\beta > 0$).

Proposition 2. *For the case of two otherwise identical carriers with $\delta_1 < \delta_2$, there exists a β^* such that the carrier 1 is willing to pay more for exclusivity for all $\beta < \beta^*$.*

If consumers are willing to trade-off handset and network quality, then the handset is worth relatively more to the lower quality carrier. Once β gets high enough, its value is sufficiently augmented by the higher quality carrier for it to be willing to pay more. This tells us that measuring whether or not consumers are willing to substitute between handset and network quality will be a determinant of a carrier's willingness to pay.

This section has established that exclusive contracts can be jointly profit maximizing depending on the relative elasticities of the two markets. The primary mechanism is through an increase in effective elasticity when setting prices, although these contracts can also encourage new entrants. When carriers are also vertically differentiated, we see that consumers' willingness to substitute between handset and network quality will affect which downstream firm values exclusivity more.

2.5   Conclusions

This paper proposes a simple motivation for exclusive contracting in the smartphone market: since consumers are more willing to substitute between downstream goods (wireless networks), an exclusive contract with an upstream firm (handset maker) can reduce price competition and lead to higher equilibrium prices. However, since the downstream goods are not in fact *perfect* substitutes, exclusivity leads to a smaller market potential, and so the question of whether or not it leads to higher joint profits of the contracting parties is an empirical question.

16

Future research directions include extending the theory model to examine the optimal length of exclusivity under some alternative assumptions, such as decreasing marginal costs or positive usage externalities. These may explain why we routinely observe shorter length exclusive contracts.

3. PRICING AND ENTRY INCENTIVES WITH EXCLUSIVE CONTRACTS: EVIDENCE FROM SMARTPHONES

3.1 Introduction

Apple launched its first ever smartphone in 2007, the iPhone, exclusively on AT&T (then Cingular) in the United States. Many handsets are released exclusively, although the Apple arrangement was notable for its 5 year term.[1] The popular press devoted much attention to the wisdom of the Apple decision, as AT&T was plagued by complaints of poor network quality with the iPhone, despite being the largest carrier in the US at the time.[2] In addition, many customers of other wireless carriers expressed interest in purchasing an iPhone, but could not do so without switching carriers. This led to political and regulatory attention being paid to exclusive contracts between handset makers and wireless networks. The Federal Communications Commission (FCC) and United States Senate have held hearings on the potentially negative impact on consumers of these arrangements.[3] The view of the major wireless carriers was that these arrangements increased welfare through greater incentives

---

[1] For example, the Palm Pre smartphones launched exclusively on Sprint, while the first touchscreen Blackberry was exclusive to Verizon and the first Blackberry Pearl exclusive to T-Mobile. Exclusive contracts are typically in the 6-12 month range.

[2] Apple Press Release "Apple Reinvents the Phone with iPhone," http://www.apple.com/pr/library/2007/01/09Apple-Reinvents-the-Phone-with-iPhone.html

[3] See Press Release, "Kerry, Wicker, Dorgan, Klobuchar Call for Increased Choice for Wireless Consumers", Sen. John Kerry, Jun 15, 2009.

for innovation, as wireless carriers have a stronger incentive to invest in new innovations for which they will be the exclusive provider.[4] The view of consumer groups was that exclusivity leads welfare losses from higher prices and fewer choices for consumers.[5] Indeed, the effect on welfare is ambiguous.[6]

This paper conducts an empirical analysis and counterfactual simulations of the forces that shaped this industry. In order to estimate the magnitudes of these competing forces, we require estimates of the price elasticities of the various handsets and wireless carriers. However, estimating demand in such a setting poses several challenges. Demand is dependent between months as this is a durable good where a consumer's current demand is a function of the consumer's current "state" (her current handset, contract status with her wireless carrier, and any switching costs that her contract imposes). A consumer's state evolves according to a known process and the consumer's history of choices. I build a choice model closely related to the Pure Characteristics Model of Berry and Pakes (2007), where random coefficients rationalize decisions and individual tastes are invariant over time. Consumers will choose between bundles every period by comparing discounted future utility flows conditional on their current state. I avoid a fully-dynamic sequential model by simplifying

---

[4] AT&T gave its "visual voicemail" feature for the iPhone as an example of such an investment. However, other carriers subsequently added this capability to their networks for handsets running Windows Mobile, Blackberry, Android, and Symbian operating systems.

[5] A specific concern was that, at the time, AT&T did not have a wireless network in several rural areas as well as the states of Vermont and Alaska. Consumers in those areas could not purchase an iPhone even if they were willing to switch carriers.

[6] This paper will not provide an estimate of the welfare effect of allowing exclusive contracts. There are two competing forces affecting welfare: higher prices in a static context, but increased entry in the dynamic context. While the effect of exclusive contracts on entry incentives can be measured, the change in entry probability is not identified, and so the latter force cannot be estimated. I can provide bounds on the latter force, but they are not informative for setting policy. For a paper that focuses on the welfare question of Apple's exclusivity, see Zhu, Liu, and Chintagunta (2011).

consumer beliefs, and argue that the simplification is supported by the data.[7] An advantage of my approach is that I avoid i.i.d. taste shocks for every product in every period.[8] I contrast my approach with a standard Logit demand model in Appendix B.3.

The econometric approach taken in this paper follows a simulated non-linear least squares (SNLLS) estimator developed by Laffont, Ossard, and Vuong (1995), which explicitly corrects for simulation bias introduced by simulation methods. This estimator is feasible for a small set of markets, but as the number of parameters grows, it becomes computationally challenging. To estimate the full model, the SNLLS estimator is nested inside an MCMC routine developed by Chernozhukov and Hong (2003), enabling the estimation of a large number of unobserved heterogeneity terms that are not recoverable via an inversion mapping, as is common in demand estimation.

This paper's contributions to the literature are an empirical investigation of exclusive contracting to the case of horizontal differentiation at both the upstream and downstream levels. Empirical applications of vertical exclusivity models are limited; for examples see Asker (2005) and Lee (2010). This paper's setting is an advantageous one in which to study the effect of downstream market power, as upstream goods are bundled one-to-one with the downstream good. The goal of the econometric analysis is to understand the the impact of consumer preferences on the observed vertical structure of an important market in the United States. The results from the econometric analysis are then used to answer three counterfactual questions: first,

---

[7] An important contribution to the dynamic discrete choice literature is Gowrisankaran and Rysman (2011), which nests a demand system within a dynamic optimization decision framework, fully internalizing for a consumer the decision to buy now or wait. An example of a prior paper which avoids dynamic programming in such a setting is Geweke and Keane (1996).

[8] As noted elsewhere, for example in Ackerberg and Rysman (2005), such taste shocks can lead to bias in elasticities in the current setting.

how much would each of the carriers have been willing to pay for exclusivity with Apple in 2007? Second, did Apple's exclusivity with AT&T increase entry incentives for Android handset makers, and if so, by how much? Finally, how much was AT&T willing to compensate Apple for each unit sale foregone due to exclusivity? Of particular interest is that the answer to the first of these questions is highly dependent on recomputing a price equilibrium. That is, if a new price equilibrium is not computed, the observed market outcome appears inefficient.

The paper proceeds as follows: Section 2 describes the industry and data I will use for the empirical analysis. Section 3 develops an econometric model of consumer choices. Section 4 discusses the results from estimation. Section 5 provides the results from counterfactual simulations. Section 6 summarizes.

3.2 Industry and Data Description

3.2.1 The United States Wireless Market

There are four nationwide wireless carriers in the United States who together control approximately 85% of the market: Verizon, AT&T, Sprint, and T-Mobile. Smaller, regional carriers account for the balance. Mobile phone penetration is high, with 95% of adults owning mobile telephones by the end of 2010. Smartphones are a fast-growing segment of mobile telephones: despite the first smartphones appearing in the 1990s, smartphones never achieved widespread consumer adoption until advances in cellular data networks and increases in the power of mobile devices led smartphones to dominate new mobile telephone purchases in 2011.[9] Smartphones differ from traditional mobile phones ("feature phones") in that they offer rich data

[9] See Nielsen Press Release "In US, Smartphones Now Majority of New Cellphone Purchases," June 30, 2011.

services such as e-mail, web browsing, photo and video capture, and multiple software applications in addition to voice features. The dominant smartphone operating systems are Apple's iOS, Google's Android, and Research in Motion's Blackberry. Of those three, Android is the only one whose owner does not control hardware as well: Google has several hardware partners that build and market smartphones, including Motorola, Samsung and HTC.

Wireless carriers purchase spectrum from the US government and construct and operate wireless networks, offering consumers various monthly packages of voice and data usage. Smartphones are typically sold on subsidized two year contracts: consumers commit to two years of a monthly plan that includes a data component in exchange for being able to purchase a smartphone at a reduced price. The subsidized price of a smartphone typically falls between $0 and $250, whereas the unsubsidized retail price is often between $500-$700. Monthly plans for smartphones range from $65 to $130, depending on the features that are included.

The fact that smartphones are sold on two-year contracts introduces the fact that the choice to buy a new handset is a dynamic one. Purchasing a handset-network bundle in the current month creates a switching cost for the next 24 months due to the early termination fee (ETF) clause common in all contracts. These fees start between $175 and $350, and decrease by $0-10 per month over the length of the contract.[10] Smartphones are subsidized by wireless carriers, so this fee prevents consumers from leaving before the subsidy has been recovered by the carrier.

---

[10] Over the time period in question, T-Mobile's ETF is $200 for the entire contract length. Verizon and AT&T are both $175 decreasing by $5 per month at the beginning of the data period but switch to $350 less $10 per month in November 2009 (Verizon) and $325 less $10 per month (AT&T) in June 2010. Sprint starts at $200 and falls by $10 per month until it reaches $50, where it remains until the end of the contract.

3.2.2 *Demand data*

I use proprietary datasets gathered by The Nielsen Company in my estimation: Nielsen conducts a monthly survey of the United States wireless telecommunications market. Between 20,000 and 25,000 individuals are contacted every month (though, not the same individuals every month) and are asked a series of individual questions including income range, age, race, gender, household size, employment, and education level. They are also asked whether or not they subscribe to mobile phone service, and if so, on which carrier and using which handset with which price plan. The geographic market of the individual is also observed, as is the time since they acquired their current handset, and whether or not they have switched carriers in the previous 12 months.[11] I have access to the survey months of November 2008 until December 2010. I omit people under 18 years of age and people who identify that their employer provided their phone to them.[12] The survey observations are assigned weights to correspond to census data. Appendix Table B.1 provides some summary statistics.

3.2.3 *Product data*

The demand dataset contains the name of the chosen handset and carrier as well as basic data on product characteristics: flags for keyboard, touch screen, smartphone, and brand. I have augmented the dataset with additional characteristics for smartphones including software operating system, processor speed, and the num-

---

[11] Unfortunately, I do not observe the previous handset-network bundle, or even the identity of the previous carrier for these individuals.

[12] Combined, these represent approximately 4% of observations.

ber of "apps" available.[13] Self-reported prices are available by device in the demand dataset, but due to the high variance in the price reported for a handset on a given carrier purchased in a given month, I omit self-reported prices for purchases that occurred more than 3 months before the survey and take the mode of reported values for a given month of purchase. Further, as some models have few reported purchases in a given month, I impose that handset prices be weakly decreasing over time.[14] Discussions with industry sources confirm that at the monthly level, prices for a given handset rarely increase. Network prices are publicly available. I choose the network price for each carrier's introductory smartphone bundle, which during this sample consists of 450 "peak" minutes (500 on T-Mobile), unlimited evening and weekend minutes, unlimited in-network calling, unlimited text message, and unlimited data. There are many combinations of features that can result in different prices, but I chose this price as many add-ons and features are the same price across networks, and so this provides a benchmark. Furthermore, these plans correspond to the modal range of base monthly prices paid by consumers in the survey data. There are other minor differences between the plan prices I use, such as different hours for what qualifies as "evening" and different definitions of "in-network calling", however I allow these differences to be absorbed by carrier fixed effects.[15]

I further augment the demand data with carrier network performance data at the market level taken from periodic "Drive Tests", where a team from Nielsen drives around a market with devices that simulate cell phones and record signal strength,

---

[13] The primary source for the added data was the database of handset characteristics maintained by the website www.phonearena.com.

[14] That is, if the median reported prices paid for a handset in months t and $t+1$ are p_t and p_{t+1}, I impose that the price in month $t + 1$ is p_t in the event that $p_{t+1} > p_t$.

[15] For example, Sprint allows free calls to any mobile number, not just other Sprint customers.

dropped calls, and other performance data of all of the available carriers in the market. This data is collected every 4-6 months for approximately 100 markets across the USA. I linearly interpolate in-between months for these metrics and match the markets to the markets identified in the demand data. The 90 markets for which I have both demand and network quality data form the basis of estimation. These 90 markets represent most of the 100 largest MSAs, covering over 190 million Americans.

I collapse all non-smartphones into a single "feature phone", available on every carrier at the same fixed price with a mean utility to be estimated. I am left with 211 handset-network bundles over the course of 26 months.[16] In terms of individual handsets, I observe 4 models of iPhones, 18 models of Blackberries, and 43 models of Android phones.

3.2.4 Data Description and Trends

There are two dominant wireless carriers in the United States: AT&T and Verizon, who each control approximately 30% of mobile customers. They are followed by Sprint (16%) and T-Mobile (11%). Network quality data appears to be highly persistent over time within a market, but exhibits significant variation across markets for all of the carriers. Figure 3.1 shows a non-parametric density plot of the rate of dropped calls across markets for each carrier in a given month and a plot of the dropped call rates within a sample market over time. Figure 3.2 provides a summary of each carrier's network quality ranks.[17] Note that for contractual reasons, there

---

[16] I perform additional data-cleaning activities, such as removing observations of T-Mobile iPhones, which were unauthorized "unlocked" models of the original iPhone.

[17] In Figures 3.1, "Carrier 0" is all other carriers besides the four national ones. There exist markets where there is no carrier beyond the four major ones, and so I omit "Carrier 0" from cross-sectional figures.

are certain pieces of data that cannot be fully labeled.[18] In the density plot, it is apparent that each of the carriers competes in markets where their network quality is "good" (few dropped calls) and others where it is "bad" (many dropped calls). However, it is also apparent that some some networks are generally "better", with their distributions concentrated to the left, and some are generally "worse", with their distributions more diffuse. In Appendix B.6, I provide evidence that network quality is exogenous, and argue that any potential bias from the endogeneity of network quality would work against my counterfactual results. The second plot shows that, in a sample market, the relative rankings of the carriers' network quality does not change over the 26 months that I use for estimation. In fact, the rates barely move at all over the 26 months. The third shows that every carrier has markets where they are ranked each of 1st, 2nd, 3rd and 4th out of the four major carriers in terms of network quality. As a comparison, Consumer Reports conducts an annual survey of 50,000 cell phone customers and publishes carrier ratings for approximately 25 metropolitan areas in every January issue.[19] For the years 2008-2011, Verizon is the highest rated carrier in their survey, although there are individual markets where other carriers are rated superior.

A key trend in this time period is the rapid adoption of smartphones. In the first month of my data, 8% of adults own a smartphone, which triples to 24% in the final month. The share of device purchases in a given month that are smartphones increases from 4% to nearly 20% during this period. In the same period, the share of adults that own any phone increases from 89% to 95%. The solid lines in Figure 3.4 shows this smartphone trend split out by income group. The mix of smartphones that

---

[18] As some summary statistics from Nielsen's research are made public, there will be occasions where firm names are included.

[19] See, for example, Consumers Union of United States (2009).

Dropped Call Distributions across Markets by Carrier

Dropped Call Distributions Within Market over Time
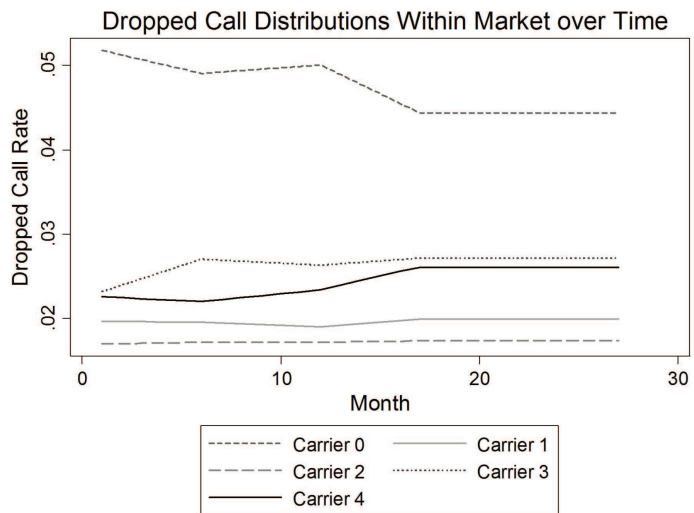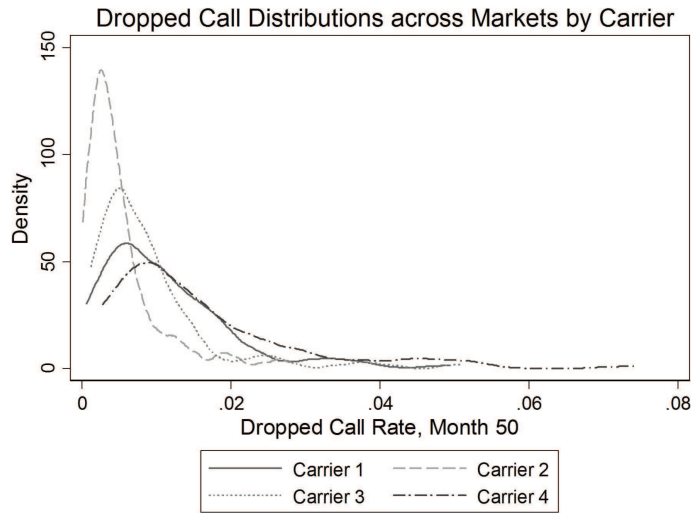
Figure 3.1: Network Quality Cross-Section vs Time Series
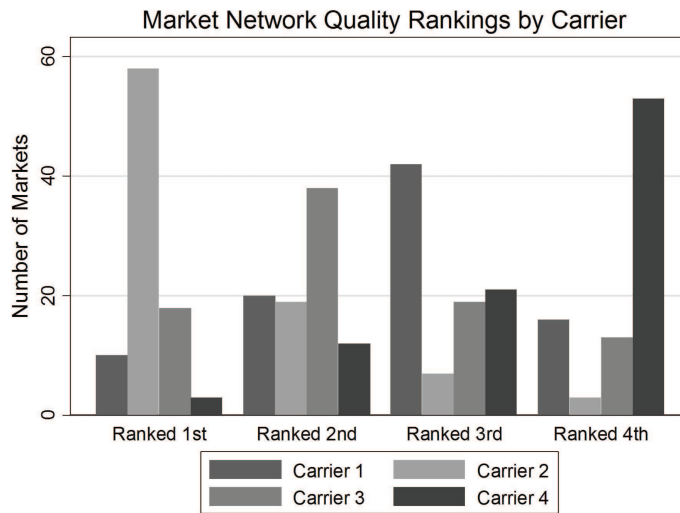
Figure 3.2: Network Quality Ranks Within Markets
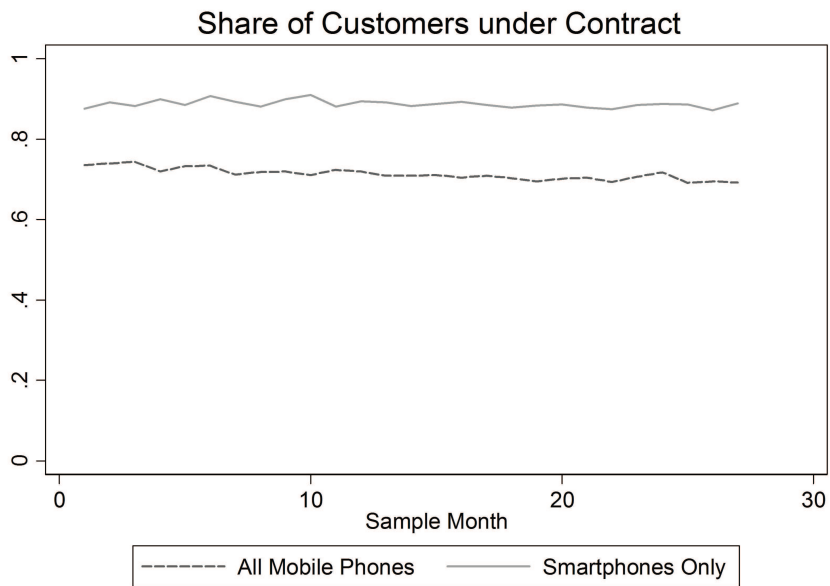
Figure 3.3: Share of Consumers on Mobile Phone Contracts

consumers own also undergoes a dramatic swing: iOS (the operating system used on iPhones) and Android see strong growth, while Blackberry's growth lags the growth of smartphones overall. The solid lines in Figure 3.5 show the share of adults that own a given type of smartphone over time. By the end of 2010, iOS and Android each control nearly 30% of smartphones.

Another interesting trend is the share of customers under contract. Figure 3.3 shows that the share of customers that are currently on a contract for their mobile phone does not change much over the sample period, even when restricted to only smartphones. Over 90% of smartphone consumers report signing a two-year contract that includes an ETF.

Additional plots of raw data are discussed with the estimation results in Section 3.4, where plots of actual versus fitted moments of the data are discussed to illustrate how well the model fits the data at the estimated parameter vector.

3.2.5 Reduced-Form Evidence

To determine whether or not consumers respond to my measure of network quality, I performed a cross-sectional regression of carrier share on dropped calls for a single month of my data, including carrier fixed effects and clustering standard errors at the market level. The first specification uses the share of minor carriers and of people with no phones as the omitted category, whereas the second specification calculates a carrier's market share as its share of the market held by the four national carriers. The results (Appendix B.2) show an effect of dropped calls significant to the 99% level, and estimate that a 1% increase in a carrier's dropped call rate translates into a decrease of market share of nearly 1%. This indicates that consumers do indeed respond to differences in network quality.

29

From the theory model, we are interested in estimating the substitutability of handsets versus wireless carriers. However, since the data are not a true panel, we cannot directly look at switching rates between different handset-network bundles. We are interested in distinguishing whether the market is composed of, say, consumers who want a Blackberry regardless of which carrier it is on, or consumers who want to be on Verizon regardless of what handset they have. Treating each market as an independent realization of preferences, we can look at the cross-section for evidence of substitution.

Consider the following: if carriers are good substitutes for one another, we would expect to see wide variance in carrier market shares across markets, relative to the variance in smartphone market shares. See Appendix Figure B.1 for plots of these shares across markets in the raw data. We can see that there does appear to be more variation in carrier market shares than in smartphone market shares across markets. However, there are obvious confounds to this: we believe that differences in network quality affect a carrier's market share, as discussed above. Similarly, since the iPhone is exclusive to AT&T, we would expect AT&T's strength in a market to affect the different smartphone market shares. Appendix Figure B.2 plots the residuals from regressions of market shares on controls. We clearly see that, controlling for relevant confounds, there is little variation in smartphone shares across markets, but large variation in carrier shares across markets, lending support to the idea that carriers are good substitutes for one another, but smartphones are poor substitutes for one another.

3.3   Empirical Model of Demand for Smartphones

Utility maximizing consumers choose every month to either consume a handset-network bundle, or to have no mobile phone (with discounted present value of utility normalized to 0). A consumer's state in a given month is what device she currently owns, the months remaining on her contract (if any), and any early termination fee (ETF) that would apply if she chose to switch to a new device or carrier. The consumer chooses between alternatives every month based on the discounted utility from each handset-network bundle.

Monthly Flow Utility   I begin with monthly flow utility: An individual i in market m receives flow utility from handset h on network n in month t that consists of a handset component, a network component, an interaction between those two, and a monthly access fee:

$$u_{imnht} \;=\; (1-\beta_t)^{(t-t_{i0})} \left[\delta_{imnt} + \delta_{iht} + \beta^c \cdot \delta_{imnt} \cdot \delta_{iht}\right] - \alpha_i \cdot p_n \qquad (3.1)$$

$$\delta_{imnt} \;=\; \beta_{in} \cdot X_{mnt} + \xi_{nm}$$

$$\delta_{iht} \;=\; \mathbf{fi}_{ih} \cdot X_{ht}$$

The term $(1-\beta_t)^{(t-t_{i0})}$ captures a deterministic rate of decay of a handset purchased in month t_{i0} over time, with the monthly decay rate β_t to be estimated. The term β^c is analogous to the one from Section 2.4, and allows consumer utility to be non-linear in the utility of the individual bundle components. Utilities from the handset and network, δ_{imnt} and δ_{iht} respectively, are modeled as projections on to the characteristics of the networks and handsets. Consumers have individual-specific tastes over network characteristics, which consist of network n's rate of dropped calls

in market m in period t.[20] Smartphone bundles also include a fixed carrier-market effect ξ_{nm} that is constant over time that captures unobserved heterogeneity in carriers across markets. Similar to network quality, handset quality depends on a vector of handset characteristics over which consumers have random and fixed coefficients: random coefficients over indicators for the Android, iOS, and Blackberry handheld operating systems, and fixed coefficients over processor speed, indicators for feature phone and smartphone, time trends in feature phone and smartphone, the log of the number of "apps" available on the handset platform, and whether or not a given device is that network's "flagship" device[21] at that time.[22] The network's monthly access fee is p_n. An individual's price sensitivity, α_i, will be modeled as

$$\alpha_i = \mathbf{Z}_i \cdot \mathbf{\beta}_\alpha + \eta_i^\alpha \tag{3.2}$$

where $\mathbf{Z}_i$ are indicators for an individual's income group,[23] β_α are fixed coefficients and η_i^α is an i.i.d. mean-zero normal draw with variance σ_η to be estimated.

---

[20] The dropped call rates used in estimation are relative to the market average. There exist markets where, for geographic reasons, all major carriers have poor quality networks, but I do not observe less adoption of mobile phones in those markets. Instead, the primary driver of differences in overall mobile phone adoption across markets is the income distributions of the markets. Conditional on owning a mobile phone, the relative shares of the carriers is heavily influenced by their relative quality, as discussed in Section 3.2.5.

[21] While I observe advertising spending by carrier and market, I do not observe it at the device level. Conversations with industry sources confirm that carriers focus their device advertising on one "flagship" device at a time. Therefore, I have identified each network's "flagship" device for the period in question, and assigned it an indicator equal to that carrier's share of advertising spending in that market and month.

[22] Additional characteristics such as GPS, wifi, memory, screen size, screen resolution, and camera resolution have also been gathered. However, trends in these are highly collinear with processor speed, and so they are not included.

[23] I use 7 income groups in total, as all groups above $100K in income have similar rates of ownership of smartphones in the dataset. Note that the mean income coefficient of the lowest income group is normalized to -1, but for the remaining groups is estimated freely.

The individual-specific random coefficients $\mathbf{fi}_i = [\beta_{in} \, \mathbf{fi}_{ih}]$ multiply the network quality characteristic and a vector of handset operating system dummies, respectively, and are distributed jointly normal according to $\mathbf{fi}_i \sim N\left(\bar{\beta}, \Sigma\right)$. All off-diagonal elements of Σ are set to 0, except those corresponding to covariances between random coefficients of the handset OS dummies and the rate of dropped calls, which are to be estimated. Note that these random coefficients are not subscripted by time period; they are persistent over time.

Discounted Flow Utility   A consumer's decision on which device to purchase is clearly a dynamic one: purchasing a device today and signing a two-year contract increases my cost of changing to a new device in the next 24 periods. However, the state space over the 24 months of a smartphone contract consists of all possible characteristics, prices, and availabilities, and so some simplification must be made to make the problem tractable. I assume that at the time of contract signing, a consumer does not expect to break her contract: she evaluates discounted utility without explicitly accounting for the option value of switching in every period between the current one and the end of her contract.[24] In the data, less than 1.4% of observations report paying termination fees in the previous 12 months. Discussions with industry sources indicate that consumers who pay such fees have often either broken their handset, rendering it useless, or are responding to a another truly unexpected event such as a relocation.[25] These are consistent with consumers not expecting to break their con-

---

[24] When estimating the model, consumers are indeed able to break their contract and switch to a different bundle. Unreported estimates from a model where consumers are not able to switch while under contract yields similar results.

[25] Given the high "retail" (unsubsidized) listed prices of handsets, if a handset is broken, it can often be less expensive to pay an ETF and purchase a new subsidized handset than to replace the previous handset.

tract at the time of signing.[26] A second challenge is how to model the continuation value at the end of a contract. I will borrow a suggestion from other dynamic discrete choice studies[27] and assume that the maximum discounted utility available from a handset-network combination in the current period is sufficient to predict future values of the maximum discounted utility available from a handset-network bundle. This is captured in the continuation value function $\gamma_{it}()$ described below.

Given the flow utility, consumer i in market m that currently owns handset h on network n with r_{it} months remaining on their contract has the following present value of utility from that handset-network combination:

$$U_{imhnt} = \sum_{m'=0}^{r_{it}-1} b^{m'} u_{imnht} + b^{r_{it}} \cdot \gamma_{it}(r_{it}) \tag{3.3}$$

In every period, a consumer will compare this value to other possible choices available to them. I use the notation $(nh)'$ to indicate an alternative handset-network bundle. A consumer's information set in the current month consists of all characteristics and prices of the products that are available. Specifically, every other handset available on every network, and the outside good of having no mobile telephone. The present value of utility from purchasing a new bundle handset-network pair in period t in market m is

$$U_{im(nh)'t} = \alpha_i \cdot \left(p_{(nh)'t} + ETF_{it} + \beta_i^s \right) + \sum_{m'=0}^{23} b^{m'} u_{im(nh)'t} + b^{24} \cdot \gamma_{it}(24) \tag{3.4}$$

The discount factor b is fixed at $0.9916 = 0.9^{(1/12)}$, giving an effective annual

---

[26] Unreported estimates of this model omitting observations who claimed to have broken contracts yields similar results to the reported results.

[27] See Gowrisankaran and Rysman (2011) and Geweke and Keane (1996)

discount rate of 10%. The term γ_{it} is a reduced-form representation of the consumer's continuation value at the end of their contract. It can be thought of as that person's value of being off contract, and will be modeled as $\gamma_{it}(x) = \theta_\gamma^x \max_{(nh)} \{U_{imnht}\}$. That is, a consumer looks at the discounted utility available from other bundles this month, and expects the maximum of that set to grow by a percentage every month.[28]

The first term in the above equation captures the cost of purchasing the handset at price $p_{(nh)'t}$, paying an early termination fee (ETF) of ETF_{it}, and paying some individual specific intrinsic switching cost β_i^s, designed to capture the cost of learning about new devices, learning how to use a new device, and transferring data. Early termination fees vary by carrier and typically decrease every month from the date of purchase until the contract expires after two years. Consumers who are off-contract in period t have $ETF_{it} = 0$. The 24-month discounting reflects the two-year length of contract.

Therefore, the consumer's decision to consume handset h on network n in a given period is captured by the inequality

$$U_{imnht} \geq U_{im(nh)'t} \quad \forall \, (nh)'$$

3.3.1 An Alternative Logit Approach

The above model is similar to the Pure Characteristics model described by Berry and Pakes (2007), which omits i.i.d. Logit draws for each possible good and opts instead for only random coefficients to rationalize tastes. If, instead, we were interested in estimating a version of this model with Logit tastes, we could indeed add i.i.d. Logit errors to each discounted flow utility U_{imnht} and directly estimate a likelihood

---

[28] The maximum of the set is selected as though the consumer were not currently on a contract, as that is the proper benchmark for modeling the value of being off-contract.

for each survey respondent. However, such a model has several drawbacks, which are discussed in Appendix B.3.

3.3.2 Estimation Approach

The approach taken to estimate the above model will be to use a simulation estimator for a small number of markets, but to nest that estimator with a Markov Chain Monte Carlo method to recover estimates for the full dataset.

The simulation estimator I use is the simulated non-linear least squares (SNLLS) estimator proposed by Laffont, Ossard, and Vuong (1995). The model described above could also be estimated using a simulated GMM estimator in the spirit of McFadden (1989) or Pakes and Pollard (1989). Given a parameter vector, the model would predict market outcomes for every market and every month given product characteristics and prices. Simulation methods could be used to integrate over the random coefficients, and the simulated moments of the model could then be matched to observed moments of the data. However, as is well-known in this literature, minimizing a naive sum-of-squares of the difference between simulated and observed moments is biased for any fixed number of simulation draws.[29] The SNLLS estimator explicitly corrects for the simulation bias in the objective function, resulting in a consistent estimator that is far less computationally demanding than alternative approaches.[30]

---

[29] See Appendix B.4 or Laffont, Ossard, and Vuong (1995) for details.

[30] An alternative approach to this problem proposed by Gourieroux and Monfort (1993) uses moment conditions of the form

$$E\left[\left(\psi_l^0 - \psi_l^{NS}\left(\theta\right)\right) \frac{\partial \psi_l^{NS}\left(\theta\right)}{\partial \theta}\right] = 0$$

where different sets of draws are used to compute the simulated moments and their derivatives, respectively, to eliminate correlation. Computing the derivative of the simulated moment is computationally costly in this setting.

A second challenge is that the unobserved heterogeneity parameters, ζ_{nm}, introduce 450 parameters that must be recovered when all 90 markets are included. In practice, no optimization routine would be able to find a global extremum over such a parameter space. The remaining parameters are not independent of the ζ_{nm} terms, complicating estimation of the full set of parameters. My approach is to use a Markov Chain Monte Carlo (MCMC) method proposed by Chernozhukov and Hong (2003) which nests the SNLLS estimator inside an MCMC framework.[31] As they show, for an estimator such as SNLLS, a Markov Chain can be constructed that shares the same distribution as the asymptotic distribution of the estimated parameter vector. Parameter estimates can be taken as the mean of the Markov Chain.

A final challenge is that this type of model faces the "initial conditions problem" (Heckman, 1981), where the process that determines a sequence of outcomes must somehow be initialized. For example, when simulating this model, most individuals already own a mobile phone in my first month of data. I cannot take this empirical distribution as given and assume that the random coefficients are distributed independently of the state observed in the first month; a given parameter vector must rationalize that initial state (as discussed in Appendix B.3). If the conditional distribution is not known, then the ideal approach is to start where there is no initial condition (Pakes, 1986). Therefore, I simulate starting 5 years before my data begins, allowing consumers to make decisions once per year in a random month, and then

---

[31] Nesting a simulation-based estimator within an MCMC approach creates a minor problem: the correction term proposed by Laffont, Ossard, and Vuong (1995) is consistent for any linear transformation of the objection function. However, the MCMC method involves an exponential transform when calculating jump probabilities to construct the chain. This results in a bias in jump probability for a fixed number of simulations, that goes to zero as the number of simulations goes to infinity. The author is aware of this issue and is currently pursuing multiple approaches to correcting for this issue. Monte Carlo experiments indicate no effect on consistency of estimates. Estimates from Specifications (1) and (2) are not affected by this issue, and comparing estimates from Specification (3) to (2) suggest it does not have a material effect on results.

up to 4 times in the final year depending on their random month.[32] The choice set in this initial period is limited to a smaller set of smartphones than truly existed, but that captures the most popular models observed in the first month of data.[33]

For practical reasons, I will first estimate the parameters of the model for a small number of markets using SNLLS, and then use these estimates as the starting point for the MCMC estimation. Estimation using the simulation estimator proceeds as follows:

1. For each of the M markets and $N = 7$ income groups, draw a set of S vectors to represent the unobservable types.

2. For each market m, determine a set of weights that, when applied to the N individuals drawn in Step 1, match the observed distributions of the N types in that market. That is, each market is expressed as a mixture of finite types of consumers. Similarly, determine weights for each market that represent their share of the national market.

3. Search over parameter vectors to minimize an objective function (discussed below). For each candidate parameter vector,

   (a) Transform a set of S draws to correspond to the random coefficients $\mathbf{fi}_i \sim N\left(\bar{\beta}, \Sigma\right)$ in accordance with the candidate parameter vector.

---

[32] I chose 5 years because 98.6% of observations in the first month of data claim to have purchased their current smartphone within 5 years; 98.0% is the average for all months.

[33] The prices and release dates for the smartphones available in this "initial period" were gathered by hand. The smartphones included are all iPhones, the Blackberry Curve, Pearl, Bold, 7200 series and 8800 series, the Motorola Q series of Windows phones, the Nokia N75 series, and a "generic" smartphone available on each carrier to capture all others. The generic "feature phone" is also included for each carrier.

(b) For all $N \cdot S$ "drawn individuals" in each of the M markets, simulate the sequence of choices for every month.[34]

(c) Calculate moments of these sequences that can be matched against observed moments of the dataset.

(d) Calculate the bias-corrected objective function.

What does a sequence of choices for a "drawn individual" look like? As an example, a sequence of choices may be that an individual in a certain market with a set of taste draws emerges from the initial period and arrives in month 1 of my data with a Blackberry on Sprint and four months remaining on contract. In months 2-7, this individual perceives greater discounted flow utility from her current device, even though her contract expired in month 5 and her handset is decaying at a monthly rate of β_t. However, in month 8, a new iPhone is released and this consumer perceives a higher level of discounted flow utility from the iPhone-AT&T bundle, even after paying for the new handset and paying an internal "switching cost".[35] This consumer buys that bundle and then remains with this bundle through month 26, as no other bundle offered enough of an increase in discounted flow utility in any of months 9-26 to overcome her contract termination fees and internal switching cost. This is a single sequence for a single drawn individual in a single market: I simulate many of such sequences for each market based on different draws of unobservables.[36] Once many sequences have been simulated, they can then be aggregated into moments

---

[34] The sequences of choices is begun 5 years prior to the start of the dataset, as discussed in Section 3.3.4.

[35] I estimate the distribution of the switching cost, β_i^s, as a normal truncated at 0 with mean μ_s and standard deviation σ_s. While this captures the implicit cost of learning a new device and transferring data between old and new devices, it may also be capturing frictions such as search costs.

[36] An important feature is that the same draw of unobservables may result in different paths in different markets, due to differences in network quality.

such as market shares or average characteristics of products (the exact moments used in estimation are discussed in Section 3.3.3).

For each moment $l = 1..L$, we want to match the simulated moment $\psi_l^{NS}(\theta)$ to its observed value in the data, ψ_l^0. The bias-corrected objective function subtracts a consistent estimate of the simulation error (discussed in Appendix B.4), resulting in

$$Q_{LNS}(\theta) = \frac{1}{L} \sum_{l=1}^{L} \left\{ \left(\psi_l^0 - \psi_l^{NS}(\theta) \right)^2 - \frac{1}{S(S-1)} \sum_{s=1}^{S} \left(\psi_{sl}^{NS}(\theta) - \psi_l^{NS}(\theta) \right)^2 \right\}$$

where $\psi_{sl}^{NS}(\theta)$ is the value of the simulated moment for a single simulation draw and $\psi_l^{NS}(\theta) = \frac{1}{S} \sum_{s=1}^{S} \psi_{sl}^{NS}(\theta)$. Thus, our consistent estimate of the parameter vector is $\theta^* = \arg\min_\theta Q_{LNS}(\theta)$.

Once the above method has recovered an estimate θ^* of the true parameter vector θ^0, the standard inference methods for simulation estimators can be used to recover confidence intervals for all parameter estimates. See Specifications (1) and (2) in the results section for estimates for limited numbers of markets using SNLLS.

The MCMC estimator uses the method developed by Chernozhukov and Hong (2003), which nests an extremum operator within an MCMC framework. The approach is to construct a quasi-posterior density over the parameter of interest according to

$$p(\theta) = \frac{e^{-Q_{LNS}(\theta)} \pi(\theta)}{\int_\Theta e^{-Q_{LNS}(\theta)} \pi(\theta) \, d\theta}$$

where Θ is a compact convex subset of $\mathbb{R}^k$ that contains θ^0, $\pi(\theta)$ is a prior probability distribution, and Q_{LNS} is the objective function from the SNLLS estimator described above. Inspection of this density reveals that it places most weight in areas of the parameter space where $Q_{LNS}(\theta)$ is small, or where the simulated model closely

matches the observed data. In order to compute an estimate of θ^0, we can construct a Markov chain whose marginal density is given by $p(\theta)$ and recover our estimates as the mean of the chain. To construct the Markov Chain, I will use the Metropolis-Hastings algorithm with quasi-posteriors suggested by Chernozhukov and Hong (2003), where from a starting value $\theta^{(0)}$, I generate a new candidate vector θ' from a conditional density $q(\theta'|\theta)$, and I update according to

$$
\theta^{(j+1)} = \begin{cases} \theta' & w.p. \, \rho\left(\theta^{(j)}, \theta'\right) \\ \theta^{(j)} & w.p. \, \left(1 - \rho\left(\theta^{(j)}, \theta'\right)\right) \end{cases}
$$

where the transition probability is given by

$$
\rho\left(\theta^{(j)}, \theta'\right) = \min\left(\frac{e^{-Q_{LNS}(\theta')} \pi(\theta') q\left(\theta^{(j)}|\theta'\right)}{e^{-Q_{LNS}(\theta^{(j)})} \pi(\theta^{(j)}) q\left(\theta'|\theta^{(j)}\right)}, 1\right)
$$

I use a standard normal for $q(\theta'|\theta)$, making the chain a random walk. That is, each candidate vector is centered at the current vector. Further, I specify a flat prior for all terms.[37] This simplifies the transition probabilities for my specification to:

$$
\rho\left(\theta^{(j)}, \theta'\right) = \min\left(\frac{e^{-Q_{LNS}(\theta')}}{e^{-Q_{LNS}(\theta^{(j)})}}, 1\right)
$$

Therefore, if a candidate vector improves the objective function, the chain moves to that point with probability 1. If a candidate vector worsens the objective function, the chain moves to that point with some positive probability that depends on the change in the objective function. Because of this, the chain spends relatively more time in the parameter space where the simulated model fits the observed data. Once

---

[37] The correlation parameters are constrained to be within the interval $[-0.9, 0.9]$. The handset decay rate is constrained to be non-negative.

the chain reaches a sufficient length, its mean $\bar{\theta}$ can be used to provide a consistent estimate of θ^0.

In summary, consumers have individual-specific taste draws for each carrier, for each of the three handset operating systems, for price sensitivity (as a function of income), for network quality, and for switching costs. These individual tastes are persistent over time. I simulate a large number of sequences of consumer decisions and match moments of the simulated model to moments of the raw data, correcting for bias introduced by simulation error. The total number of parameters to estimate is 34, plus the 5 carrier-market fixed effects per market, for a total of 485 parameters when using all data.

3.3.3 Identification and Moments

Given that this is a non-linear model, there is not a one-to-one mapping between moments and the parameters that they identify. Nonetheless, it is useful to consider what sources of variation in the data are likely to be influencing different parameters. Network monthly access prices do not change over this period, and so identification of preferences for networks comes primarily from cross-sectional variation in the quality and market share of each network, controlling for each market's income distribution. Prices and characteristics of handsets are changing significantly over time but are the same across markets, and so the time-series variation in these are identifying preferences for handsets, as well as parameters relating to switching costs and the handset decay rate. The variation in ownership rates of feature phones and smartphones between income groups identifies differences in price sensitivities between income groups.

A common concern when estimating tastes for a bundle of two goods (a handset

and a network in this case) is confounding correlation of tastes with complementarity between the elements of the bundle.[38] In this setting, these separate elements are identified by the cross-section variance in network quality. If, for example, tastes for Blackberries and network quality are correlated, I would expect to see the share of consumers with Blackberries roughly similar across markets, but that consumers sort into the higher quality carriers in each market. If instead, the two elements of the bundle are complements, then I would expect a carrier's share of consumers with smartphones to increase across markets as its network quality increases.

The moments used in estimation are the following for each of the 26 months of data: The first set of moments are market-level shares of each carrier for all phones, and for smartphones only. These influence parameters of tastes for network quality, as well as the variance of network tastes and the market-carrier fixed effects. The second set of moments are national-level shares of each smartphone operating system and average characteristics by smartphone operating system (including network quality). These moments drive the taste parameters for the handset operating systems and characteristics, as well as the correlation parameters between handset types and network quality. The third set is the rate of smartphones purchase. This informs structural parameters such as switching costs and the rate of handset decay. The fourth set of moments are the share of ownership of smartphones, and any phone, by income group. These help isolate price coefficients as well as mean utilities and time trends. The total number of moments being estimated is 24,076 when all 90 markets are included in estimation. When estimating for a single market, the number of moments is 936.

---

[38] See Gentzkow (2007) for an analysis of this issue in the context of online newspapers versus print newspapers.

Estimation of the SNLLS parameters was done using a simulated annealing algorithm in Matlab, using "mex" files to simulate consumer choices and calculate moments and the objective function. I use Halton Sequence draws to improve coverage and reduce spurious correlation. The distribution of random coefficients for dropped calls and for switching costs are truncated at 0, so that no one may get positive utility from dropping calls or switching devices. The MCMC chain constructed has a total length of 100,000 after a burn-in of 10,000 draws. I group the parameters for the Metropolis-Hastings algorithm into the following groups: (1) price coefficients, (2) characteristics, (3) the ζ_{nm}, and (4) all remaining parameters. The variance of the draws for each parameter group is adjusted after every 100 draws per group to maintain an acceptance rate as close as possible to 0.5.

3.4 *Estimation Results and Discussion*

See Figures starting at 3.4 for plots of fitted moments (dashed lines) versus actual moments (solid lines). Parameter estimates start in Table 3.1 for four different specifications: the first specification is estimated using SNLLS on a single market with 2,800 effective draws from the unobserved parameter vector. The second is estimated using SNLLS on 6 markets using 4,200 effective draws. The third uses the identical setup as Specification 2 but switches the estimation approach to MCMC, to show that the SNLLS and MCMC approaches produce similar results. The fourth is estimated using MCMC on 90 markets with 18,900 effective draws.[39] Figure 3.7 gives examples of

---

[39] In all tables, a dash (-) for a standard error indicates that the parameter was fixed in the given specification. Any parameters listed with a μ or σ are indicating that the estimated parameters are means and standard deviations of random normal variables, respectively.
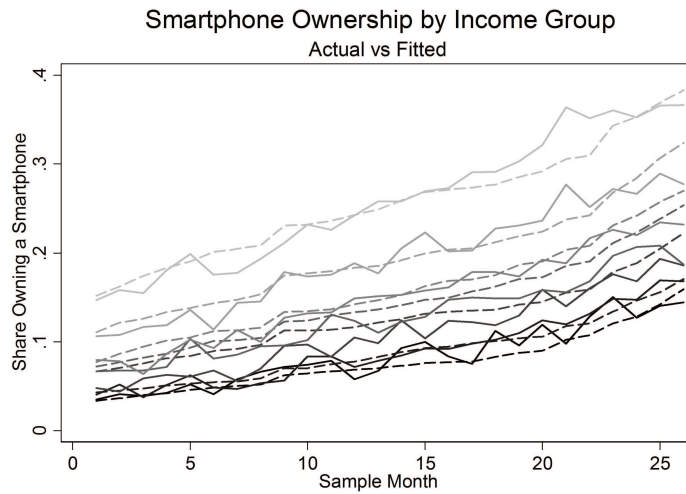
Figure 3.4: Fitted vs Actual: Smartphone Ownership by Income

Note: Solid lines represent actual data. Dotted lines are fitted. Different shades represent different income groups; top (lightest) line is incomes of 100K+, decreasing in order to lowest line representing incomes of $15K or less. Ordering reflects ordering in Table 3.1. Based on results from Specification 4.

the MCMC process. The first panel shows the acceptance rate for the first parameter group. The second panel shows the movements of a single parameter, the price coefficient mean for group with income of $100K+. The vertical black bars indicate the transition between the burn-in period and the period used in calculating estimates. As can be seen, the process appears to settle into a stationary distribution before the end of the burn-in period.

Of greater interest than the individual parameter estimates are estimates of price elasticities for each carrier's monthly access price for smartphones. Table 3.7 shows estimates of price elasticities for each carrier's monthly smartphone access price at the observed monthly smartphone access prices and handset contracts in column 1. The second column of Table 3.7 shows estimates of these same price elasticities if Apple's

45

Shares of Smartphone Operating Systems
Installed Base
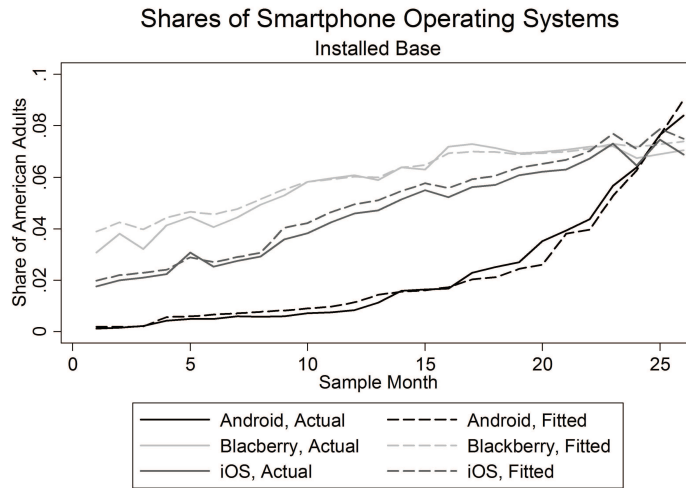
Figure 3.5: Fitted vs Actual: Handset O/S Shares

Note: Solid lines represent actual data. Dotted lines are fitted. Based on results from Specification 4.

Smartphone Adoption by Carrier
Actual vs Fitted
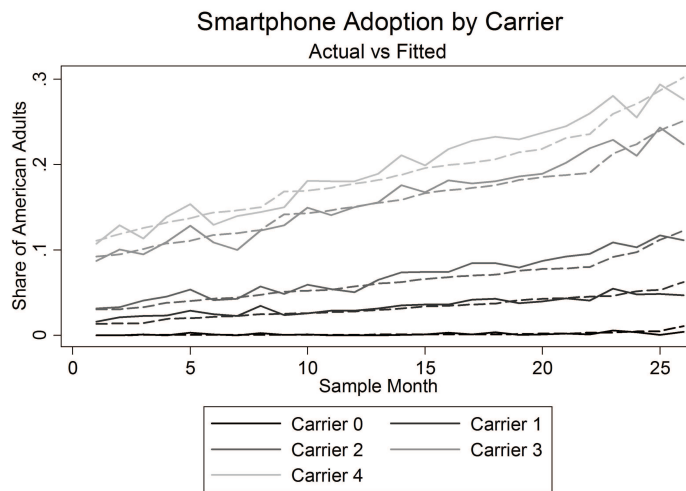
Figure 3.6: Fitted vs Actual: Carrier Shares of Smartphones

Note: This graph stacks the share of American adults with a smartphone on a given carrier, showing actual (solid line) versus fitted (dashed line), using estimated from Specification 4.

Table 3.1: Price Coefficient Estimates

Income Group	Specification			
	(1)	(2)	(3)	(4)
¡$15K	-1	-1	-1	-1
	-	-	-	-
$15-25K	-0.97614	-0.9759	-0.9871	-0.9758
	(0.0692)	(0.0464)	(0.0406)	(0.0325)
$25-35K	-0.9345	-0.9340	-0.9556	-0.9340
	(0.0993)	(0.0382)	(0.0241)	(0.0118)
$35K-50K	-0.9143	-0.9138	-0.9337	-0.9139
	(0.0416)	(0.0389)	(0.0341)	(0.0252)
$50-75K	-0.8978	-0.8973	-0.8966	-0.8973
	(0.1033)	(0.0489)	(0.0416)	(0.0275)
$75-100K	-0.8579	-0.8578	-0.8589	-0.8578
	(0.2419)	(0.0936)	(0.1223)	(0.0439)
$100K+	-0.80818	-0.8129	-0.8296	-0.8231
	(0.4109)	(0.0217)	(0.0330)	(0.0207)
Standard Deviation	0.1583	0.1578	0.1595	0.1623
	(0.0472)	(0.0319)	(0.0388)	(0.0176)
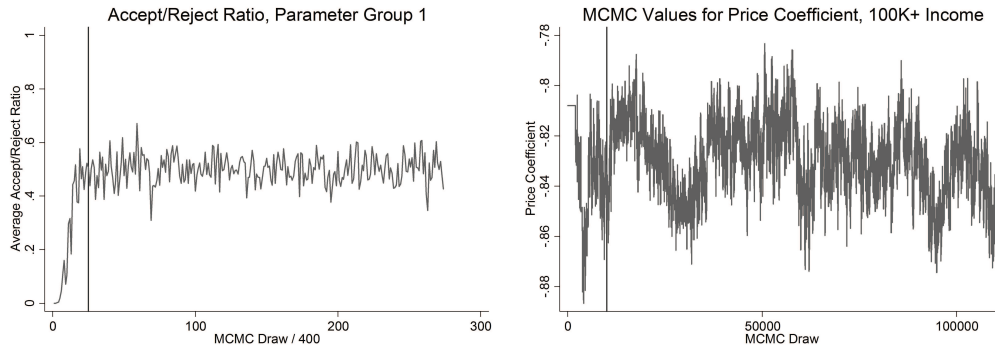
Figure 3.7: MCMC Convergence Charts

Note: Vertical bars indicate the end of the "burn in" period. Based on results from Specification 4.

Table 3.2: Handset Parameter Estimates

	Specification			
	(1)	(2)	(3)	(4)
Android μ	-9.8975	-9.9737	-9.923	-10.225
	(6.918)	(4.353)	(4.541)	(2.232)
Android σ	7.1402	7.1697	6.8054	7.1763
	(7.164)	(5.291)	(4.541)	(2.366)
iOS μ	-3.9514	-3.9692	-3.842	-3.969
	(8.662)	(7.930)	(6.837)	(2.461)
iOS σ	5.8701	5.887	6.063	5.897
	(2.214)	(5.024)	(4.622)	(1.752)
Blackberry μ	-21.517	-22.162	-22.046	-22.204
	(11.449)	(8.702)	(10.377)	(5.297)
Blackberry σ	18.721	18.647	18.602	18.581
	(13.201)	(5.092)	(3.836)	(3.771)
Log(Apps)	1.9621	1.9792	1.7743	1.9796
	(0.553)	(0.848)	(0.750)	(0.320)
Processor Speed (GHz)	1.1777	1.1857	1.2754	1.1773
	(0.778)	(0.821)	(0.684)	(0.727)
Flagship Device	0.7843	0.7898	0.7113	0.7896
	(0.485)	(0.352)	(0.257)	(0.0674)

Table 3.3: Network Parameter Estimates

	Specification			
	(1)	(2)	(3)	(4)
Voice Mean Utility	44.793	44.221	44.855	44.551
	(4.457)	(1.972)	(1.378)	(0.570)
Voice Time Trend	5.3892	5.6967	5.8096	5.7028
	(1.964)	(1.591)	(1.735)	(0.671)
Data Time Trend	2.0452	1.9618	1.9479	1.9796
	(0.775)	(0.646)	(0.781)	(0.201)
Dropped Calls μ	-20	-24.024	-24.084	-24.056
	-	(14.265)	(8.163)	(1.310)
Dropped Calls σ	20	17.077	16.896	17.072
	-	(9.002)	(10.816)	(4.524)

Time trends are based on log(month), where month begins at 1 in the "initial period", 5 years before the data begin.

Table 3.4: Carrier Parameter Estimates

	Specification			
	(1)	(2)	(3)	(4)
Carrier 0 σ	-	0.2314	0.2196	0.2314
("all other" carriers)	-	(0.2613)	(0.2648)	(0.0725)
Carrier 1 σ	0.2969	0.3086	0.2968	0.3015
	(0.2495)	(0.3114)	(0.2967)	(0.0964)
Carrier 2 σ	0.4108	0.4138	0.4319	0.4139
	(0.1613)	(0.2409)	(0.2512)	(0.0941)
Carrier 3 σ	0.5300	0.5376	0.5449	0.5341
	(0.4688)	(0.2062)	(0.2552)	(0.117)
Carrier 4 σ	0.3317	0.3373	0.3708	0.3668
	(0.1328)	(0.3325)	(0.4352)	(0.1964)

Table 3.5: Correlation Coefficient Estimates

	Specification			
Dropped Call Correlation with	(1)	(2)	(3)	(4)
Android	-	-0.125	-0.129	-0.127
	-	(0.1858)	(0.1602)	(0.0661)
iOS	-	-0.0582	0.0745	-0.0519
	-	(0.1645)	(0.1786)	(0.1412)
Blackberry	-	-0.394	-0.102	-0.2951
	-	(0.2187)	(0.0404)	(0.0812)

Note: Since dropped calls are considered "bad", a *negative* correlation between handset taste and dropped calls indicates that people who prefer that handset also *dislike* dropped calls.

Table 3.6: Remaining Parameter Estimates

	Specification			
	(1)	(2)	(3)	(4)
Switching Cost μ	73.613	84.273	83.997	84.396
	(22.331)	(24.850)	(18.367)	(7.774)
Switching Cost σ	93.942	97.013	96.616	97.074
	(30.029)	(21.826)	(32.143)	(14.218)
Handset Decay Rate (β_t)	0.00229	0.00232	0.0018	0.007305
	(0.00443)	(0.00571)	(0.00169)	(0.00231)
Continuation Value (θ_γ)	1.0023	1.0035	1.0052	1.0049
	(0.00585)	(0.00581)	(0.00214)	(0.00199)
Handset-Network Complementarity (β^c)	-0.00155	-0.00156	-0.00211	-0.00185
	(0.00079)	(0.00155)	(0.00144)	(0.000861)

Table 3.7: Elasticity Estimates

	Specification	
	(1)	(2)
Carrier A	-0.9360	-1.0111
Carrier B	-1.0046	-0.2333
Carrier C	-0.6078	-13.465
Carrier D	-1.2881	-1.3346

Estimates are at parameter vector estimated in specification (4) above. Estimates are price elasticity of demand of monthly smartphone access price over time dataset time period.

iPhone had been available on all carriers, holding all prices fixed. These are measured as the change in total quantity of monthly subscribers over the entire sample period, for a change in the monthly access price.

3.4.1 Discussion

There are a number of trends to highlight in the parameter estimates. First, many parameters are estimated more sharply as the number of draws and number of markets used increase. A large number of parameters are not significant when using only a single market (Specification 1). This is to be expected, as characteristics such as the network quality vary across markets much more than they do over time. Therefore, we would expect parameters such as the distribution of tastes for network quality and the correlation parameters to be poorly estimated with few markets. Second, the MCMC method provides similar results to the SNLLS for the overlapping specifications. This comparison provides a consistency check that the MCMC method provides an equivalent approach to the SNLLS method. The parameters that are least similar between the two are parameters such as the correlations, which are poorly estimated in general with a small number of markets. Third, note that the 6 markets used for Specifications 2 and 3 are the six largest markets in the sample. These appear

to be a selected group, as some parameters show large swings when moving to the 90 markets, particularly those that are identified by cross-market variation.

Looking at the parameter estimates themselves, we see that the price coefficients are sharply estimated and are decreasing in magnitude as income increases. All characteristic coefficients have the expected sign. Other parameters of interest show that there is weak evidence of consumers substituting between handset and network quality, and that the most significant correlation between handset and network tastes is with the Blackberry, where consumers who have positive taste for Blackberries also have stronger disutility from dropped calls. The estimated distribution of switching costs has a mean of approximately $80, but a large standard deviation. Handsets decay at a rate of approximately 1% per month.

The plots of fitted moments versus actuals based on Specification 4 show that at the estimated parameter vector, the simulated model fits most of the data well (starting at Figure 3.4).

Of most interest to the theoretical discussion are the estimated carrier price elasticities in Table 3.7. Computing price elasticities is complicated by the fact that they are "state dependent": a consumer's tastes, characteristics, and current product choice affect their response to a carrier's price change. Consequently, the table presents estimates of elasticities based on the total quantity of smartphone-months over the data sample period. In Column 1, we see that at the observed prices, Carrier C appears to face demand that is inelastic. However, this confounds the elasticity with respect to the devices that carrier offers. Moving to Column 2 where the iPhone is available on all carriers, we see that Carrier C, the iPhone's exclusive carrier, sets a monthly access price that seems far too high should the iPhone be non-exclusive. In addition, we see that Carrier B, who has the highest quality network, is pricing far too low.

Table 3.8: Counterfactual Simulation 1: Carrier Willingness to Pay

	Carrier Willingness to Pay		
Scenario	AT&T	Verizon	Sprint
Prices Fixed	$14.12B	$20.54B	$3.02B
Prices Recomputed	$21.81B	$3.20B	$6.82B
Prices Fixed, $\beta^c = 0$	$19.85B	$32.12B	$5.66B
Prices Recomputed, $\beta^c = 0$	$23.90B	$5.43B	$9.86B

Note: Table shows each carrier's maximum willingness to pay for exclusivity with Apple, defined as the profit difference between exclusivity and the worst case of rival exclusivity.

This is evidence that the prices we observe in the market are unlikely to have been optimal had Apple not signed an exclusive contract. This will be a driving factor in counterfactual simulation.

3.5   Counterfactuals

3.5.1   Willingness to Pay for Exclusivity

This counterfactual examines, ex-ante, which of the national wireless carriers had the most to gain from an exclusive contract with Apple in 2007. Of most interest are the values for AT&T and Verizon, as these are the carriers that were rumored at the time to be in discussions with Apple. Prices for the iPhone device are fixed at their values from AT&T regardless of the carrier, but monthly access prices are allowed to re-adjust where indicated in Table 3.8. The scenarios determine the net change in monthly fee income from November 2008 until December 2010 for all carriers when the exclusive carrier is Sprint, Verizon, and AT&T. The willingness to pay is defined as the total profit with exclusivity less the profits from AT&T having exclusivity (for AT&T, it is compared to Verizon having exclusivity).[40]

If prices are held fixed (first column of Table 3.8), we see that Verizon has the high-

---

[40] Verizon is considered to include Alltel, even though that merger was announced in June of 2008.

est willingness to pay, as they are able to attract a large number of subscribers when offering the iPhone. However, the theory motivation presented earlier indicated that the primary driver of exclusivity being optimal is the change in price equilibrium. In order to determine a new price equilibrium, I numerically estimate the price elasticity of demand for each carrier at the estimated parameter vector, and use this to recover an estimate of each carrier's marginal cost. Then, taking that marginal cost as given, I re-assign the iPhone devices to other carriers, and starting at the observed prices, iterate best responses for each carrier until a new equilibrium is found.[41] I then determine the change in profits at the new equilibrium.

Once prices are allowed to adjust (second column of Table 3.8), we see that AT&T has a significantly higher willingness to pay. This is due to the fact that AT&T's equilibrium price without exclusivity is lower than Verizon's. Verizon enjoys less elastic demand, and so has less harm from rival exclusivity than AT&T.

The final two columns change the estimate of β^c, the degree to which consumers are willing to trade-off between handset and network quality, to 0. The purpose of this is to determine how much this substitution is affecting willingness to pay. Since setting this parameter to 0 effectively increases utility from all handsets, we cannot compare the values to those of the first two columns. However, we still observe the large reversal in willingness to pay once prices float. This is evidence that consumers' substituting handset and network quality is far less of a factor in determining willingness to pay than the shift in price equilibrium.

---

[41] I cannot prove that there is a unique equilibrium.

3.5.2 *Effect of Apple Exclusivity on Android Entry Incentives*

This counterfactual considers the expected profits of the Open Handset Alliance[42] had Apple instead chosen to be available on more than one carrier. The scenario compares the variable profits from handsets earned from sales of Android units between November 2008 and December 2010 under the assumptions that the iPhone had initially launched on AT&T and Verizon, or on all four national carriers. All characteristics are held constant at their observed values.[43] Estimates are reported for the case where network prices are held constant at their observed values, and also when they are allowed to float to new optimal prices. Marginal contribution per handset is assumed to be $139, and is the average handset subsidy paid by the three largest wireless carriers in Q4 2010.

As can be seen in Table 3.9, the exclusivity between Apple and AT&T created a significant opportunity for the Android handset manufacturers. Consistent with the theory model, had Apple not chosen to be exclusive, expected profits for Android handset makers would have been lower by approximately $850M if the iPhone had also launched on Verizon, and nearly $1B if the iPhone had launched on all carriers. In the interest of comparing magnitudes, the 2010 net profit of HTC, one of the most successful Android handset makers, was $1.3B.[44] Therefore, this is a sizable change in incentives.[45] We can conclude that the existence of exclusive contracts creates a

---

[42] The "Android Consortium", a consortium of 84 companies that includes 22 handset manufacturers, among them Motorola, Samsung, and HTC.

[43] The most obvious characteristic that may change would be the number of "apps" available on Android, as we might expect this to be a function of the installed base of Android phones. This leads to a more conservative estimate of the number of lost sales. Future work will examine this more closely.

[44] HTC Corporation 2010 Annual Report.

[45] Furthermore, it is a conservative estimate. In addition to the issue mentioned in the previous footnote, this does not take into account changes in subsidies or handset prices. It is not feasible to recompute a new handset price equilibrium given the number of prices this would involve (every

Table 3.9: Counterfactual Simulation 2: Android Entry Incentives

| | Apple enters on: | |
	Two Carriers	Four Carriers
Android Entry Incentive	-$875.2M	-$961.4M

Note: Table shows projected change in contribution margin for Android handset makers from Apple entering on multiple carriers instead of being exclusive to AT&T.

significant incentive for entry in this setting.

3.5.3   *Apple Exclusivity vs Non-Exclusivity*

This counterfactual combines the results from the previous two to answer the question of how much Apple could have been compensated for the sales it lost due to exclusivity. Looking at AT&T's willingness to pay calculated above, the amount that AT&T would have willing to pay Apple for every iPhone that they *could have sold*, had they not been exclusive, is $148.33. This is based on AT&T's willingness to pay as computed in Counterfactual 1, and the number of handsets Apple could have sold under non-exclusivity in Counterfactual 2. As a comparison, Apple's 2010 net income for the entire firm was $14B, and the firm sold 40M iPhones worldwide.[46] If half of the current year's profits are from current iPhone unit sales, we get $175 profit per unit, which is comparable to what AT&T would have willing to compensate Apple for unit sales foregone due to exclusivity. Without more details on Apple's per-unit profit level, it is not possible to conclusively state that exclusivity was optimal, but this calculation shows that AT&T's willingness to pay was comparable to what Apple is likely able to earn per iPhone sold.[47]

---

handset, every month).

[46] Apple Corporation 2010 Annual Report

[47] Some may argue that the relevant comparison is with the case where Google's Android does not enter, as Apple may not have anticipated Android's 2008 entry into the market. However, Google had

3.6 Conclusions

This paper proposes a simple motivation for exclusive contracting in the smart-phone market: since consumers are more willing to substitute between downstream goods (wireless networks), an exclusive contract with an upstream firm (handset maker) can reduce price competition and lead to higher equilibrium prices. However, since the downstream goods are not in fact *perfect* substitutes, exclusivity leads to a smaller market potential, and so the question of whether or not it leads to higher joint profits of the contracting parties is an empirical question.

An econometric analysis of this market shows that consumers are far more price sensitive with respect to wireless networks than handsets, and so exclusivity may be a profit-maximizing strategy. Counterfactual simulations show that AT&T was indeed willing to sufficiently compensate Apple for the smaller market potential caused by exclusivity, and that this exclusive contract significantly increased the entry incentives of rival smartphones, such as those running Google's Android operating system.

Future research directions include extending the theory model to examine the optimal length of exclusivity under some alternative assumptions, such as decreasing marginal costs or positive usage externalities. These may explain why we observe shorter length exclusive contracts and why Apple renegotiated its exclusivity with AT&T before the end of the 5-year term.

---

purchased the software developer responsible for Android in 2005, and so it is reasonable to assume that Apple anticipated such an entry.

4. COMPETITION AND IDEOLOGICAL DIVERSITY: HISTORICAL EVIDENCE FROM US NEWSPAPERS[1]

4.1 Introduction

Economists have long been concerned with the optimal amount of product diversity in the marketplace (Dixit and Stiglitz 1977, Mankiw and Whinston 1986). In the context of the news media, product diversity matters not only for the usual reasons of consumer and producer surplus, but also because it may contribute to the competitiveness of the marketplace of ideas and hence of the political process (Becker 1958, Downs 1957). Thus, "the [First] Amendment rests on the assumption that the widest possible dissemination of information from diverse and antagonistic sources is essential to the welfare of the public".[2]

Three main policy instruments have been directed at increasing ideological diversity in media markets: explicit subsidies, relaxation of antitrust rules, and limits on joint ownership. Federal, state, and local governments in the United States subsidized newspapers in the nineteenth and early twentieth centuries, and many European governments continue to do so today, with the explicit goal of maintaining diversity (Murschetz, 1998). The Newspaper Preservation Act of 1970 allowed competing newspapers to jointly set advertising and circulation prices in an effort to prevent

---

[1] Joint with Matthew Gentzkow and Jesse M. Shaprio, University of Chicago Booth School of Business

[2] United States Supreme Court, "Associated Press v. United States," 326 US 1.

second and third papers from exiting. The Act states its goal as "maintaining a newspaper press editorially and reportorially independent and competitive in all parts of the United States." The Federal Communications Commission has long regulated US media ownership "on the theory that diversification of mass media ownership serves the public interest by promoting diversity of program and service viewpoints" (FCC, 2010).

In this paper, we study the economic forces that determine equilibrium ideological diversity in newspaper markets. We formulate an equilibrium model of entry and product positioning, with competition for both consumers and advertisers. We show descriptive evidence consistent with the model's core predictions, and estimate the model using data on the circulation and affiliations of US daily newspapers in 1924. We then use the estimated model to decompose the incentives that promote diversity and evaluate the impact of the public policies discussed above.

Studying newspapers in a historical context affords several advantages that offset the intrinsic disadvantage of moving further away from contemporary policy settings. First, during the time period that we study it was common for newspapers to declare explicit political affiliations (Gentzkow, Glaeser, and Goldin 2006, Hamilton 2006). A newspaper's affiliation serves as a good proxy for the ideological tilt of the newspaper's content (Gentzkow, Shapiro, and Sinkinson, 2011), so the presence of explicit affiliations alleviates the challenge of measuring ideology that confronts studies of modern news media (Groseclose and Milyo 2005, Gentzkow and Shapiro 2010). Second, during the period we study there were a large number of local markets in the US with multiple competing daily newspapers. Although many media remain fiercely competitive today, few afford researchers a large cross-section of experiments that can be used to study competitive interactions.

Partisanship emerges as an important determinant of newspaper demand. Within a metropolitan area, an increase of 10 percentage points in the proportion of a town's votes going to Republicans increases the relative circulation of Republican papers in the town by 10 percent. Adding a second Republican paper to a town with one Republican and one Democratic newspaper reduces the relative circulation of the existing Republican paper by 4 percent. These findings survive flexible controls for the quality of the newspaper and for the town's overall taste for news.

Such patterns in demand should induce newspapers to choose affiliations commensurate with the ideology of the local market, and to choose affiliations different from those of local competitors. Both patterns are present in our data. A 10 percentage point increase in a market's fraction Republican increases the probability that an entering newspaper chooses a Republican affiliation by 23 percentage points. Controlling for the fraction Republican, adding an additional Republican incumbent reduces an entering paper's likelihood of choosing a Republican affiliation by 15 percentage points.

Our economic model embeds the multiple-discrete-choice demand framework of Gentzkow (2007) in a sequential entry game in the spirit of Bresnahan and Reiss (1991) and Mazzeo (2002). In the model, firms first decide whether to enter the market, then choose either Republican or Democratic affiliation, taking into account household demand, the responses of other entering firms, and the effect of affiliation choice on subscription and advertising prices. The model allows households to exhibit a preference for newspapers whose ideology matches their own, and to regard newspapers with the same political affiliation as more substitutable than newspapers with different affiliations. Our model of advertising demand builds on the recent two-sided markets literature in allowing advertisers to place advertisements in mul-

tiple newspapers and to value "single-homing" and "multi-homing" consumers differently (Armstrong 2002, Ambrus and Reisinger 2006, Anderson, ystein Foros, and Kind 2011).

A crucial identification issue arises from unobserved heterogeneity in household ideology. Such heterogeneity will cause the choices of firms within a given market to be positively correlated, biasing downward estimates of the incentive to differentiate. It will also bias demand estimates, for similar reasons. We address this issue by allowing explicitly for unobserved cross-market variation in household ideology, which is identified by correlation of choices across markets that are close enough to share similar characteristics but far enough apart that their newspapers do not compete. We assume in the spirit of Murphy and Topel (1990) and Altonji, Elder, and Taber (2005) that the spatial correlation in unobservable dimensions of ideology matches that of observable measures. Experiments with specifications that ignore unobservable heterogeneity show that even qualitative conclusions of the model are sensitive to the quality of the econometrician's observable proxies for ideology, whereas conclusions from a model that allows for unobservable heterogeneity are robust.

We find that competition plays a critical role in driving ideological diversity. Newspapers with the same affiliation are better substitutes than newspapers of different affiliations, creating a strong incentive to differentiate. This effect is enhanced by competition in both circulation and advertising prices. Were entering newspapers to ignore the presence of competitors in choosing their affiliations, the number of "diverse" news markets with at least one paper affiliated with each political party would decline by almost half.

We use the model to simulate the effects of various public policies that are often motivated by a desire to maintain diverse news markets. Antitrust leniency, in the

60

form of joint operating agreements that permit pricing and advertising collusion, decreases the incentive to differentiate, but increases entry. On net, this policy increases the share of households living in markets with diverse papers from 28 percent to 41 percent. Joint operating agreements also increase consumer welfare, both through increased entry and through lower prices that result from the increased attractiveness of consumers to advertisers. Although advertisers lose surplus under joint operating agreements, total social welfare rises. Newspaper subsidies such as US postal subsidies or direct press subsidies (such as those in many European countries) affect diversity mainly through their impact on the number of newspapers.

Our work builds on recent empirical models of entry and product positioning with explicit demand systems (Reiss and Spiller 1989, Einav 2007and 2010, Draganska, Mazzeo, and Seim 2009, Seim and Waldfogel 2010, Fan 2010). Like Fan 2010, we study a news market with both subscription and advertising sides. Our model differs from past work in allowing unobserved shocks at both the firm-level and the market-level. We show that market-level heterogeneity is important in our setting, and that properly accounting for it has a significant impact on our substantive results.

Our paper also contributes to the literature on two-sided markets. Consistent with recent theoretical work (Armstrong 2002, Ambrus and Reisinger 2006, Anderson, ystein Foros, and Kind 2011), we find that the nature of advertising competition depends crucially on the extent to which consumers read multiple newspapers. We show that this force, in turn, has an important effect on firms' incentive to differentiate from their competitors. Along with Fan 2010, ours is among the first empirical studies to estimate a micro-founded model of advertising competition. In this sense, we extend past empirical work by Rysman (2004), Kaiser and Wright (2006), Wilbur (2008), Argentesi and Filistrucchi (2007), Chandra and Collard-Wexler (2009), Sweet-

ing (2010), and others.

Substantively, our paper is most closely related to research on the incentives that shape the political orientation of the news media. Gentzkow and Shapiro (2010) use a similar framework to study ideological positioning of US newspapers in recent years. Because few modern markets have more than one newspaper, however, they cannot address the impact of competition. Other related work studies the way content relates to electoral cycles (Puglisi, 2011), economic conditions (Larcinese, Puglisi, and Snyder, 2007), political scandals (Puglisi and Snyder, 2008), and government influence (Durante and Knight Forthcoming, Qian and Yanagizawa 2010), without explicitly modeling the role of competition. The Chiang (2010) study of US newspapers is the closest to ours in investigating equilibrium positioning of newspapers in multi-paper markets. Chiang (2010) uses household-level data to test the predictions of a variant of the Mullainathan and Shleifer (2005) model, and finds that ideologically extreme households in multi-paper markets are more likely to read a newspaper than those in single-paper markets.

Like Chiang (2010) and Gentzkow and Shapiro (2010), we focus on the commercial, rather than political, incentives of news outlets. Commercial considerations likely dominated political incentives at the time of our study (Baldasty, 1992). In other work, we show that newspapers' affiliations exert, on average, at most a small effect on electoral outcomes (Gentzkow, Shapiro, and Sinkinson, 2011), and that incumbent parties exert at most a limited influence on newspapers' political affiliations (Gentzkow, Petek, Shapiro, and Sinkinson, 2011). We note, however, that Petrova (2009) provides evidence that political patronage influenced newspaper affiliations in the late 1800s.

The remainder of the paper is organized as follows. Section 4.2 introduces the

historical data that forms the basis of our analysis. Section 4.3 discusses the historical context for our data. Section 4.4 lays out our economic model. Section 4.5 presents descriptive evidence on the determinants of newspaper demand and affiliations. Section 4.6 details our econometric assumptions and explains how we implement our estimator. Section 4.7 discusses model identification. Section 4.8 presents estimates. Section 4.9 presents counterfactual simulations. Section 4.10 concludes.

4.2 Data

4.2.1 Cross-section of Daily Newspaper Markets

We define the universe of potential daily newspaper markets to be all cities with populations between 3,000 and 100,000 and at least one *weekly* newspaper as of 1924. Data on the universe of cities and their populations comes from the 1924 N. W. Ayer & Son's *American Newspaper Annual.* In appendix C.1 we present an analysis of the sensitivity of our findings to tightening the population bounds for the sample and to excluding cities close to very large cities.

We take data on daily newspapers from the US Newspaper Panel introduced in Gentzkow, Shapiro, and Sinkinson (2011). The data are drawn from annual directories of US newspapers from 1869 and from every presidential year from 1872 to 1924, inclusive. In each year, we extract the name, city, political affiliation, and subscription price of every English-language daily newspaper. We match newspapers across years on the basis of their title, city, and time of day. Gentzkow, Shapiro, and Sinkinson (2011) provide details on data collection and validation of data quality.

We define a time-constant measure of affiliation for each newspaper, where papers are classified as Republican if they ever declare a Republican affiliation and Democratic if they ever declare a Democratic affiliation. In the handful of cases where a

newspaper declares a Republican affiliation in one year and a Democratic affiliation in another, we use the majority affiliation. We exclude 142 newspapers whose only affiliation is Independent and 36 newspapers that never declare an affiliation of any kind from our sample. In appendix C.1 we present results for the subsample of markets that do not contain an independent newspaper in 1924 and the subsample that do not contain an unaffiliated newspaper in 1924.

For each market in our universe with two or more daily newspapers, we define the order of entry by the order in which the papers appear in the US Newspaper Panel. When necessary we break ties randomly.

We match markets to Census place definitions in 1990 and match each Census place to the county containing the largest share of the place's population in 1990. We use the Census place-county match to combine city level newspaper data with county level voting data from various sources, as in Gentzkow, Shapiro, and Sinkinson (2011). Our main measure of consumer ideology is the average share of the two-party presidential vote going to Republicans over the period 1868 to 1928. We exclude a small number of markets for which we cannot identify the presidential vote share. In appendix C.1 we present results excluding markets in the South, where the Democrats were dominant.

Table 4.1 presents summary statistics for our cross-section of markets. Our sample includes 1910 markets, 950 of which have at least one daily newspaper, and 338 of which have more than one daily newspaper. Population is highly correlated with the number of newspapers. In total there are 1338 newspapers in the sample, of which 57 percent are Republican. Overall, 54 percent of multi-paper markets are ideologically diverse in the sense of having at least one Republican and at least one Democratic newspaper. In the average market, Republican and Democratic presidential candi-

64

Number of Newspapers	0	1	2	3+	All
Mean population	5944	10688	24049	36832	10943
Share of newspapers that are Republican		.60	.50	.68	.57
Share of multi-paper markets that are diverse			.53	.61	.54
Republican vote share					
Mean	.52	.51	.50	.55	.51
Standard deviation	.15	.15	.12	.09	.15
Number of markets	960	612	297	41	1910
Number of newspapers	0	612	594	132	1338

Notes: Data are from cross-section of markets. Diverse markets are those with at least one Republican and at least one Democratic newspaper. Republican vote share is the average Republican share of the two-party vote in presidential elections from 1868-1928.

dates tend to get a similar number of votes, but there is substantial cross-market variation in the vote share.

4.2.2 Town-level Circulation Data

We assemble a separate cross-section of towns that are close enough to newspaper markets that newspapers circulate in them, but that are not the headquarters of any daily newspaper themselves. These "hinterland" towns will be the basis of our demand analysis. Data on circulation by town comes from the 1924 Audit Bureau of Circulations (ABC) Auditor's Reports of individual newspapers. In most cases the audits cover a twelve-month period ending in 1924; in some cases the examination period is shorter or ends in 1923. We obtained the reports on microfilm from ABC. A document imaging firm scanned the microfilm, and a data entry firm converted the scanned reports to machine readable text. ABC audit reports are a standard source for newspaper circulation data, but as far as we know this is the first effort to digitize a full report from the early twentieth century.

From each audit report we extract the paper's name, location, and circulation in each town that receives "25 or more copies daily through carriers, dealers, agents, and mail." We sum circulation by town across multiple editions of the same paper and average circulation by town across multiple audit reports (if more than one edition or audit report is available).

We match newspapers in the ABC data to papers in the US Newspaper Panel using the paper's name and location. We construct a cross-section of towns with at least one matching circulating newspaper. We exclude from our sample any town that is itself the headquarters of a daily newspaper. For computational reasons, we exclude 52 towns with more than 10 newspapers available. Not all newspapers are represented in the ABC data. In appendix C.1 we present results excluding towns for which newspapers headquartered nearby are not represented in the data. We also present results from a sample that includes towns that are themselves the headquarters of a daily newspaper.

We match towns to 1990 Census place codes using town and state name, and we use place codes to match towns to counties. We exclude towns that we cannot successfully match to Census geographies, and a small number for which we do not have county presidential voting data.

Table 4.2 presents summary statistics for the towns in our sample. Our sample includes 12198 towns, in 8052 of which more than one daily newspaper circulates. Overall, 53 percent of multi-paper towns are ideologically diverse in the sense of having at least one Republican and at least one Democratic newspaper available.

Table 4.2: Summary Statistics: Towns with Circulation Data

Number of Circulating Newspapers	1	2	3+	All
Mean population	450	389	580	477
Share of newspapers that are Republican	.52	.54	.57	.55
Share of multi-paper towns that are diverse		.38	.67	.53
Republican vote share				
Mean	.49	.51	.54	.51
Standard deviation	.16	.16	.15	.16
Number of towns	4146	3737	4315	12198
Number of newspaper-towns	4146	7474	17221	28841

Notes: Data are from towns with circulation data. Diverse towns are those with at least one Republican and at least one Democratic newspaper. Republican vote share is the average Republican share of the two-party vote in presidential elections from 1868-1928.

4.2.3 Cost and Revenue Data

We obtain 1927 balance sheet data on 94 anonymous newspapers from the Inland Daily Press Association (Yewdall, 1928). We match each record in the US Newspaper Panel to the record in the balance sheet data with the closest circulation value. Performing this match allows us to estimate cost and revenue components for each newspaper in the panel.

We compute the marginal cost of each newspaper as the annual per-copy cost of printing and distribution, including paper and ink costs and mailing and delivery costs. We also compute the annual per-copy advertising revenue of each newspaper. Finally, we compute the annual per-copy circulation revenue of each newspaper (revenue from subscriptions and single-copy sales).

4.3 Background on Newspaper Partisanship

The median newspaper in our 1924 cross-section entered its market prior to 1896. At that time it was common for newspapers to choose explicit partisan affiliations

(Gentzkow, Glaeser, and Goldin 2006, Hamilton 2006). The practice faded over time: by the mid-twentieth century it was rare for entering newspapers to declare an explicit affiliation.

A newspaper's affiliation played a clear role in determining its likely appeal to different readers. For example, in 1868, the Democratic *Detroit Free Press* announced, "The *Free Press* alone in this State is able to combine a Democratic point of view of our state politics and local issues with those of national importance" (Kaplan, 2002, 23). Similarly, in 1872, the Republican *Detroit Post* declared as its mission "To meet the demands of the Republicans of Michigan and to advance their cause" (Kaplan, 2002, 22).

Anecdotal evidence supports the view that newspapers' affiliations depended on those of competing newspapers in the same market. James E. Scripps declared in 1879 that "As a rule, there is never a field for a second paper of precisely the same characteristics as one already in existence. A Democratic paper may be established where there is already a Republican; or vice versa; an afternoon paper where there is only a morning; a cheap paper where there is only a high-priced one; but I think I can safely affirm that an attempt to supplant an existing newspaper...of exactly the same character has never succeeded" (quoted in Hamilton 2006, 47). Through the early twentieth century, James' brother, E.W. Scripps, exploited the nominal independence of his newspaper chain to adapt editorial content to market conditions, emphasizing Republican ideas in markets with established Democratic newspapers, and Democratic ideas when Republicans were entrenched (Baldasty, 1999, 139).

In Gentzkow, Shapiro, and Sinkinson (2011) we report the results of a quantitative content analysis of newspapers that uses the mentions of Republican and Democratic presidential candidates as a proxy for the political orientation of a newspaper's con-

tent. The analysis shows that the partisanship of a newspaper's content is strongly related to its political affiliation and is not strongly related to the political orientation of voters in the market once we condition on political affiliation. Moreover, for newspapers that switched from being partisan to independent, historical political affiliation remains a strong predictor of the newspaper's content. As we argue in more detail in Gentzkow, Shapiro, and Sinkinson (2011), these findings support measuring political affiliations as permanent and binary (Republican/Democrat).

As noted above, we exclude papers that never declare a Republican or Democratic affiliation from our sample. The set of completely unaffiliated papers includes many specialized commercial newspapers (e.g., mining industry news) that can plausibly be treated as separable in demand from affiliated newspapers. The set of papers that only declare Independent affiliation is more likely to include competitors to those we study. A content analysis of Independent newspapers (not shown) shows that Independent papers' orientation is, if anything, even more related to local market ideology than that of affiliated papers, though the two relationships are not statistically distinguishable. This suggests that it may be reasonable to think of Independent papers as having unreported affiliations. In appendix C.1 we present results for the subsample of markets that do not contain an independent newspaper in 1924.

4.4   Model

4.4.1   Overview

We consider a cross-section of markets, each of which has a large number of potential entrants. For now we consider the game that occurs in a particular market; we introduce market subscripts when we turn to estimation below.

We index the J newspapers that choose to enter in equilibrium by $j \in \{1, ..., J\}$.

Each entering newspaper chooses a political affiliation $\tau_j \in \{R, D\}$, a circulation price $p_j \geq 0$, and a pair of advertising prices described below. We denote the vectors of types and circulation prices chosen by all entering newspapers by τ and p respectively. The market has S households indexed by i, each of which has a political affiliation $\theta_i \in \{R, D\}$. We denote the share of households with $\theta_i = R$ by ρ and assume that ρ is common knowledge to all potential entrants.

The profits of entering newspaper j are given by

$$\pi_j = S\left[(p_j + a_j - MC)\, q_j - \xi_j\left(\tau_j\right)\right] - \kappa \tag{4.1}$$

where a_j is newspaper j's advertising revenue per copy sold, MC is a marginal cost common to all newspapers, q_j is the share of households purchasing newspaper j, $\xi_j\left(\tau_j\right)$ is an affiliation-specific variable cost, and κ is a fixed cost.

The game proceeds in five stages. First, the potential entrants choose sequentially whether or not to enter. Second, the newspapers that have entered observe their own ξ_j and sequentially choose their political affiliations. Third, newspapers simultaneously choose their circulation prices. Fourth, newspapers simultaneously choose their advertising prices. Finally, households make purchase decisions and profits are realized. At the end of each stage, all newspapers' choices are observable to all other firms. The only elements of a given newspaper j's profit function that are private information are the variable costs $\xi_j\left(\tau_j\right)$. We describe the stages from last to first. At the end of this section, we describe a separate (unmodeled) process that determines which newspapers are available in each hinterland town.

Our demand specification follows Gentzkow (2007). In the model consumers can consume any bundle of the J available newspapers, or no newspapers at all. For consumers in newspaper markets, we assume that the available newspapers are those headquartered in the market.

Households differ in the utility they get from consuming a given bundle. Let $\mathbb{B} = \mathcal{P}\left(\{1, ..., J\}\right)$ denote the set of all possible bundles of newspapers, with $B \in \mathbb{B}$ denoting a generic bundle. Household i's utility from bundle B is given by

$$U_i\left(B\right) = u\left(\theta_i, B\right) + \epsilon_i\left(B\right) \tag{4.2}$$

where $\epsilon_i\left(B\right)$ is a type-I extreme value error i.i.d. across households and bundles. The function $u\left(\theta, B\right)$ denotes the mean utility from consuming bundle B for households with affiliation θ.

We define mean utilities $u\left(\theta, B\right)$ as follows. Let $k\left(B\right)$ denote the number of distinct two-newspaper subsets of bundle B such that the two newspapers have the same political affiliation. We write:

$$u\left(\theta, B\right) = \sum_{j \in B} \left(\underline{\beta}\mathbf{1}_{\theta \neq \tau_j} + \overline{\beta}\mathbf{1}_{\theta = \tau_j} - \alpha p_j\right) - k\left(B\right)\Gamma \tag{4.3}$$

where $\mathbf{1}$ denotes the indicator function. The mean utility from consuming no newspapers is normalized to $u\left(\theta, \varnothing\right) = 0$. A household receives per-newspaper utility $\overline{\beta}$ for each newspaper in the bundle that has the same affiliation as the household, and per-newspaper utility $\underline{\beta}$ for each newspaper that has a different affiliation. The household's utility is diminished by an amount Γ for every pair of newspapers with the same affiliation and by α for every dollar spent. Consistent with existing empirical

evidence (Kaiser and Song, 2009), we assume that consumer utility does not depend directly on the quantity of advertising.

Each household chooses its utility-maximizing bundle. Let $q_j(\theta)$ denote the share of households of type θ who purchase newspaper j. Then

$$q_j(\theta) = \frac{\sum_{\{B \in \mathbb{B}: j \in B\}} exp(u(\theta, B))}{\sum_{B' \in \mathbb{B}} exp(u(\theta, B'))}. \tag{4.4}$$

The market-wide share of households purchasing newspaper j is then

$$q_j = \rho q_j(R) + (1 - \rho) q_j(D). \tag{4.5}$$

4.4.3 Advertising Prices

There exists a unit mass of potential advertisers. If a household sees its advertisement in k different newspapers, an advertiser receives a benefit of $a_h + (k-1) a_l$, where $0 \le a_l \le a_h$. If $a_l = a_h$, an advertiser's payoff is proportional to the number of impressions its advertising receives. If $a_l < a_h$, the model exhibits diminishing returns beyond the first impression. If $a_l = 0$, an advertiser cares only about whether or not a household is reached by its advertisement. The difference between a_l and a_h therefore captures the extent of diminishing returns in advertising impressions.

After circulation prices are chosen, each newspaper simultaneously declares an advertising price. After advertising prices are posted, each advertiser simultaneously decides whether or not to advertise in each newspaper.

Denote the share of firm $j's$ readers who read only newspaper j by ψ_j. In any pure strategy equilibrium, all advertisers advertise in all newspapers. Newspaper $j's$

advertising revenue per reader, a_j, is given by

$$a_j = a_h \psi_j + a_l \left(1 - \psi_j\right).$$ (4.6)

Each newspaper charges advertisers for the incremental value of the impressions it delivers (Armstrong 2002, Anderson, ystein Foros, and Kind 2011). Because of diminishing returns in the value of impressions, a newspaper's advertising revenue per reader is increasing in the fraction of its readers who read it exclusively.

4.4.4 *Circulation Prices*

All newspapers that have entered the market choose prices simultaneously, having observed the set of entrants and their affiliations τ. An equilibrium of this game is a vector of prices p^* such that each element p_j^* satisfies:

$$p_j^* \in \operatorname*{argmax}_{p_j} \left(p_j + a_j \left(p_j, p_{\sim j}\right) - MC\right) q_j \left(p_j, p_{\sim j}\right).$$ (4.7)

Here we represent explicitly the fact that demand (and hence advertising prices) depend on the prices charged by the newspapers. We write $p_{\sim j}$ to denote the vector of newspaper j's competitors' prices.

We denote by $v_j = \left(p_j + a_j - MC\right) q_j$ the equilibrium variable profit of newspaper j net of the affiliation-specific variable cost $\xi_j(\tau_j)$.

4.4.5 *Political Affiliations*

Entering newspapers choose their affiliations sequentially in order of their indices j. Each newspaper observes the affiliation choices of preceding newspapers. Let τ_{j-} and τ_{j+} denote vectors of affiliations of newspapers with indices less than and greater

than j, respectively. Newspaper j's expected variable profit upon choosing τ_j is:

$$\overline{v}_j\left(\tau_j, \tau_{j-}\right) = E_{\tau_{j+}} v_j\left(\tau_j^-, \tau_j, \tau_j^+\right). \tag{4.8}$$

We make explicit here the dependence of a newspaper's variable profit on its own affiliation choice and the choices of the other newspapers. The expectation is taken with respect to newspaper j's conjecture about the affiliation choices of the newspapers that follow it.

The equilibrium is a vector of choices τ^* such that each τ_j^* satisfies:

$$\tau_j^* \in \operatorname*{argmax}_{\tau_j \in \{R,D\}} \overline{v}_j\left(\tau_j, \tau_{j-}\right) - \xi_j\left(\tau_j\right). \tag{4.9}$$

The shock $\xi_j\left(\tau_j\right)$ is private information and is revealed to newspaper j after it chooses to enter and before it chooses its affiliation. We assume that $\xi_j\left(\tau_j\right)/\sigma_\xi$ is distributed type I extreme value i.i.d. across newspapers and affiliations, where $\sigma_\xi > 0$ is a constant that scales the variability in the cost shocks.

Given past affiliations τ_{j-}, newspaper j chooses affiliation τ_j with probability

$$P_j\left(\tau_j, \tau_{j-}\right) = \frac{\exp\left[\frac{1}{\sigma_\xi}\overline{v}_j\left(\tau_j, \tau_{j-}\right)\right]}{\sum_{\tau \in \{R,D\}} \exp\left[\frac{1}{\sigma_\xi}\overline{v}_j\left(\tau, \tau_{j-}\right)\right]}. \tag{4.10}$$

Given realized variable profits $v_j - \xi_j\left(\tau_j\right)$ for each newspaper j, there is a unique equilibrium vector of affiliation choices that can be characterized by backward induction. The last newspaper J takes as given the affiliation choices of all preceding newspapers, so it knows $v_J\left(\tau_J, \tau_{J-}\right) - \xi_J\left(\tau_J\right)$ with certainty. Newspaper $J-1$ integrates over the distribution of $\xi_J\left(\tau_J\right)$ to assess newspaper $J's$ probability of choosing each possible affiliation, as a function of newspaper $J-1's$ affiliation choice and that

74

of all preceding newspapers. And so on.

4.4.6 Entry

After entry, indices are assigned at random and cost shocks $\xi_j(\tau_j)$ are realized. Let $P(\tau, J)$ denote the equilibrium probability of affiliation vector τ as of the entry stage (i.e., before cost shocks are realized). Then the expected variable profit of each entering firm as of the entry stage is

$$V(J) = \frac{1}{J} \sum_{j=1}^{J} \sum_{\tau} \left[P(\tau, J) E\left((v_j - \xi_j(\tau_j)) \,|\tau \right) \right]. \tag{4.11}$$

Here, the conditional expectation $E\left((v_j - \xi_j(\tau_j)) \,|\tau \right)$ reflects the fact that newspaper j chooses its affiliation after observing its cost shocks $\xi_j(\tau_j)$.

We define an equilibrium of the entry game to be a number of newspapers J^* such that, in expectation, entering newspapers are profitable but a marginal entrant would not be. That is,

$$V(J^*) \geq \frac{\kappa}{S} > V(J^* + 1). \tag{4.12}$$

If $V(1) < \frac{\kappa}{S}$ then it is an equilibrium for no newspapers to enter.

4.4.7 Circulation in the Hinterland

Each newspaper may be available for circulation in one or more hinterland towns. These towns' contribution to total circulation is small, so we ignore them in the entry and affiliation choices that we model above. However, we use data on town-level circulation to identify the parameters of our demand model.

The decision about whether to make a newspaper available in a given town is made based on expected variable profit, and any fixed and variable costs of trans-

portation.

Expected variable profit depends on expected circulation. We assume that demand for newspapers in towns follows the same structure assumed above for markets. Therefore circulation depends on the share of households in the town that are Republican ρ, the number of households S, and the number and affiliations of available newspapers in the town.

In equilibrium, the number and affiliations of the available newspapers will therefore be a function of ρ, S, and (possibly town-specific) fixed and variable costs of transportation.

4.5 Descriptive Evidence

Before turning to formal estimation, we present descriptive evidence from our data on the economic forces captured in the model.

4.5.1 Partisanship and Newspaper Circulation

In our model household utility depends on (i) the match between the newspaper's type and the household's type and (ii) the presence of substitute newspapers in the household's consumption bundle.

As table 4.3 illustrates, both factors play a significant role in driving observed demand. The table presents OLS regressions of the difference in mean log circulation between Republican and Democratic newspapers on measures of household ideology and/or the presence of substitutes. Specification (1) includes only household ideology, specification (2) includes only counts of available newspapers, and specification (3) includes both. Specification (4) adds county fixed effects to control carefully for household characteristics. Given the construction of the dependent measure, coeffi-

Demand for Partisanship

Dependent variable: Average log(circulation) of Republican papers - Average log(circulation) of Democratic papers

	(1)	(2)	(3)	(4)
Republican vote share	0.8634		0.9702	
	(0.1913)		(0.1984)	
Number of Republican papers		-0.0217	-0.0395	-0.1330
		(0.0135)	(0.0137)	(0.0210)
Number of Democratic papers		0.0054	0.0159	0.1109
		(0.0145)	(0.0147)	(0.0262)
County fixed effects?				X
R2	0.0104	0.0009	0.0133	0.5685
Number of counties	1215	1215	1215	1215
Number of towns	4287	4287	4287	4287

Notes: Data are from demand estimation sample. Models are OLS regressions. The dependent variable in each column is the difference in mean log circulation of Republican and Democrat newspapers. Republican vote share is the average Republican share of the two-party vote in presidential elections from 1868-1928. Standard errors in parentheses are clustered at the county level.

cients can be interpreted as the marginal effect of a given variable on the circulation of Republican papers relative to Democratic papers.

The greater is the Republican share of households in a town, the greater will be the relative circulation of Republican newspapers. However, having more Republican newspapers available will tend to depress the circulation of the average Republican paper due to substitution effects. Because Republican newspapers are more likely to be available in towns with more Republican households, these two effects tend to work in opposite directions. Therefore, we expect that specification (1) understates the effect of household ideology and specification (2) understates the importance of substitutes. Specification (3) shows that, as expected, both effects are estimated to be larger when the regression includes measures of both household ideology and the presence of substitutes. Specification (4) shows that using county fixed effects to

control carefully for household characteristics further increases the estimated substitution effects.

The estimated relationships in specification (3) are economically significant. Increasing the fraction Republican among voters by 10 percentage points increases the relative circulation of Republican papers by 10 percent. Adding a second Republican paper to a market with one Republican and one Democratic newspaper reduces the relative circulation of the existing Republican paper by 4 percent.

The evidence in the data that household ideology and the presence of substitutes influence newspaper demand is quite robust. In the online appendix, we present evidence from a specification that uses a fixed-effects strategy similar to that of Gentzkow and Shapiro (2010) to isolate the effect of these forces from variation in newspaper quality and the quality of the outside option. We find similar qualitative conclusions to those we report here.

4.5.2 Determinants of Newspapers' Affiliation Choices

Given that households demand own-type newspapers and that same-type papers are more substitutable, we would expect that newspaper affiliation would respond both to household ideology and to market structure.

Table 4.4 shows that these expectations are borne out in our data. The table presents OLS regressions of a dummy for whether a newspaper chooses a Republican affiliation on measures of household ideology and incumbent affiliations. Specification (1) includes only household ideology, specification (2) includes only incumbent affiliations, and specification (3) includes both. Specification (4) adds market fixed effects, identifying the effect of incumbents solely from the order of entry.

The more Republican are the households in a market, the more likely is an en-

Table 4.4: Determinants of Newspaper Affiliation

Dependent variable: Dummy for newspaper choosing Republican affiliation

	(1)	(2)	(3)	(4)
Republican vote share	2.1824		2.3350	
	(0.0557)		(0.0611)	
Number of Republican papers		-0.0145	-0.1483	-0.3931
		(0.0310)	(0.0332)	(0.0698)
Number of Democratic papers		-0.0168	0.1308	0.5260
		(0.0380)	(0.0304)	(0.0755)
Market fixed effects?				X
R2	0.3561	0.0003	0.3816	0.8384
Number of markets	950	950	950	950
Number of newspapers	1338	1338	1338	1338

Notes: Data are from supply estimation sample. Models are OLS regressions. Republican vote share is the average Republican share of the two-party vote in presidential elections from 1868-1928. The number of Republican and Democratic paper variables report the number of incumbent papers of each type at the time each paper enters. Standard errors in parentheses are clustered at the market level.

tering paper to choose a Republican affiliation. However, facing a Republican incumbent reduces the likelihood that an entering paper affiliates with the Republican party. Because Republican incumbents are more likely in markets with more Republican households, these two effects tend to work in opposite directions. Therefore, we expect that specification (1) understates the effect of household ideology, and specification (2) understates the effect of incumbent affiliation. Specification (3) shows that, as expected, both effects are estimated to be larger when the regression includes measures of both household ideology and incumbent affiliations. Specification (4) shows that the effect of incumbent affiliations survives controls for marked fixed effects.

The effects we estimate in specification (3) are economically significant. A 10 percentage point increase in the fraction Republican among households increases the likelihood of a Republican affiliation by 23 percentage points. Having a Republican incumbent instead of a Democratic incumbent reduces the likelihood of a Republican

affiliation by 28 percentage points.

4.6   Estimation

In this section we lay out the stochastic assumptions that we impose in estimation. We estimate the model in two steps. The first step estimates the demand system via maximum likelihood. The second step estimates the remaining parameters via maximum likelihood, taking as given the demand parameters from the first step. We refer to the second step as the "supply" model for convenience, although both demand and supply parameters ultimately influence firm conduct. We present stochastic assumptions first for the supply model, then for the demand model.

4.6.1   Supply Model

Index markets by $m \in \{1, ..., M\}$. Our identification strategy will exploit spatial correlation of ρ_m across markets. We assume that each market is paired with a single neighboring market and that ρ_m is correlated within pairs but independent across pairs. We define a mapping $n : \{1, ..., M\} \rightarrow \{1, ..., M/2\}$ such that markets m and m' are in the same pair if and only if $n(m) = n(m')$. We take as given an observable estimate Z_m of the share of households that are Republican.

We assume that ρ_m has an unobservable component that varies at both the pair and market level. Let $\delta_{n(m)}$ be a pair-specific unobservable distributed i.i.d. normally across pairs with mean μ_δ and variance σ_δ^2. Let η_m be a market-specific unobservable distributed i.i.d. normally across markets with mean 0 and variance σ_η^2. The distributions of $\delta_{n(m)}$ and η_m are assumed to be independent of one another and of Z_m. We assume that

$$\rho_m = \text{logit}^{-1}\left(\text{logit}(Z_m) + \delta_{n(m)} + \eta_m\right). \tag{4.13}$$

The logit transformation ensures that $\rho_m \in [0,1]$. We impose the following restriction on the covariance structure of the unobservables:

$$R \equiv \frac{Cov\left(\text{logit}\left(Z_m\right), \text{logit}\left(Z_{m'}\right)\right)}{Var\left(\text{logit}\left(Z_m\right)\right)} = \frac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_\eta^2} \tag{4.14}$$

for any m and m' such that $n\left(m\right) = n\left(m'\right)$.

Let $G\left(x|S_m\right)$ denote the CDF of fixed costs per household $\frac{\kappa_m}{S_m}$ conditional on population S_m. We assume that

$$G\left(x|S_m\right) = \text{logit}\left(\frac{x - \mu_\kappa^0 - \mu_\kappa^1 log\left(S_m\right)}{\sigma_\kappa}\right), \tag{4.15}$$

i.e. that $\frac{\kappa_m}{S_m}$ is distributed logistic with mean $\mu_\kappa^0 + \mu_\kappa^1 log\left(S_m\right)$ and dispersion parameter σ_κ. In appendix C.1 we present results from a specification that adds greater flexibility to the dependence of $\frac{\kappa_m}{S_m}$ on S_m.

The observed data consist of the affiliation vector τ_m, the number of firms J_m, the population S_m, and the observed share Republican Z_m. We treat the affiliation vector τ_m and the exact number of firms J_m as unobserved in any market with $J_m > \bar{J}$ for a cutoff value $\bar{J}$. (Note that we do not incorporate information on observed prices in the likelihood function.)

To derive the likelihood of the data, begin by supposing the econometrician can also observe the true share Republican among households, ρ_m. In this case, the like-

lihood of a given market m, which we can denote by $L_m(\rho_m)$, can be written as

$$
L_m(\rho_m) = \begin{cases} (1 - G(V(J_m + 1, \rho_m)|S_m)) P(\tau_m, J_m, \rho_m) & \text{if } J_m = 0 \\ (G(V(J_m, \rho_m)|S_m) - G(V(J_m + 1, \rho_m)|S_m)) P(\tau_m, J_m, \rho_m) & \text{if } J_m \in \{1, ..., \bar{J}\} \\ G(V(\bar{J}, \rho_m)|S_m) & \text{if } J_m > \bar{J} \end{cases}
$$

$$(4.16)$$

Here we make explicit that both $V()$ and $P()$ depend on ρ_m.

In fact the econometrician does not observe ρ_m. Therefore the likelihood L_n for a given pair n of markets m and m' integrates over the joint distribution of ρ_m and $\rho_{m'}$:

$$
L_n = \int_{\rho_m} \int_{\rho_{m'}} L_m(\rho_m) L_{m'}(\rho_{m'}) \, dF(\rho_m, \rho_{m'}|Z_m, Z_{m'}) \, d\rho_m d\rho_{m'} \tag{4.17}
$$

where $F()$ is the conditional CDF of the joint distribution of ρ_m and $\rho_{m'}$. The log likelihood of the data is then the sum of the log of L_n across all pairs.

4.6.2   Demand Model

Index hinterland towns in the ABC data with at least one newspaper of each affiliation available by $t \in \{1, ..., T\}$. We group towns into pairs and assume that the distribution of ρ_t conditional on Z_t follows the same parametric form as it does for markets m. We do not constrain the parameters of the distribution of ρ_t to equal those for ρ_m. (That is, we allow the analogues of σ_δ, σ_η, μ_δ, and R to differ.)

As with markets, let J_t denote the number of newspapers available in town t and τ_t denote their affiliations. Let S_t denote town population. We treat J_t as nonstochastic in estimation. In appendix C.1 we show that our results are robust to modeling J_t as a random variable whose distribution depends on S_t and ρ_t.

To address the endogeneity of τ_t with respect to ρ_t, we allow that the share of

Republican papers in a town is a stochastic function of ρ_t. We assume that:

$$\Pr\left(\tau_{jt} = R\right) = \text{logit}^{-1}\left(\mu_\rho^0 + \mu_\rho^1 \text{logit}\left(\rho_t\right)\right) \tag{4.18}$$

independently across newspapers j in town t. We think of this as an econometric approximation to the economic process by which news agents and other decision-makers decide which newspapers to transport to which towns, a process that we do not model explicitly. The approximation we use allows for a positive correlation between the (unobserved) share of readers who are Republican and the observed share of available newspapers that are Republican. In appendix C.1 we present results from a specification that adds greater flexibility to the dependence of $\Pr\left(\tau_{jt} = R\right)$ on ρ_t.

Let $\hat{Q}_{jt}$ denote the measured circulation of newspaper j in town t. We assume that

$$\hat{Q}_{jt} = q_{jt} S_t \zeta_{jt} \tag{4.19}$$

where q_{jt} is the share of households in town t who purchase newspaper j and ζ_{jt} is measurement error with $ln\zeta_{jt} \sim N\left(0, \sigma_\zeta\right)$ i.i.d. across newspapers and towns.

In each town, the econometrician is assumed to observe only the difference in mean log circulation between Republican and Democratic newspapers. We impose this restriction because it intrinsically scales out variation in population, which is likely to be poorly measured and therefore a significant source of heterogeneity in observed circulation.

To derive the likelihood function, suppose that the econometrician observes ρ_t in

each town. Then the likelihood $L_t(\rho_t)$ of a given town t is:

$$L_t(\rho_t) = \frac{1}{\tilde{\sigma}_{\zeta t}} \phi \left(\frac{\frac{\sum_j I_{\tau_{jt}=R} ln(\hat{Q}_{jt}/q_j)}{\sum_j I_{\tau_{jt}=R}} - \frac{\sum_j I_{\tau_{jt}=D} ln(\hat{Q}_{jt}/q_j)}{\sum_j I_{\tau_{jt}=D}}}{\tilde{\sigma}_{\zeta t}} \right) \Pr(\tau_t|\rho_t, J_t) \qquad (4.20)$$

where ϕ denotes the standard normal PDF and

$$\tilde{\sigma}_{\zeta t} = \sigma_\zeta \sqrt{\frac{1}{\sum_j I_{\tau_{jt}=R}} + \frac{1}{\sum_j I_{\tau_{jt}=D}}}. \qquad (4.21)$$

In fact the econometrician does not observe ρ_t. Therefore the likelihood L_n for a given pair n of towns t and t' integrates over the joint distribution of ρ_t and $\rho_{t'}$ conditional on Z_t and $Z_{t'}$:

$$L_n = \int_{\rho_t} \int_{\rho_{t'}} L_t(\rho_t) L_{t'}(\rho_{t'}) dF(\rho_t, \rho_{t'}|Z_t, Z_{t'}) d\rho_t d\rho_{t'} \qquad (4.22)$$

where $F()$ is the conditional CDF of the joint distribution of ρ_t and $\rho_{t'}$. The log likelihood of the data is then the sum of the log of L_n across all pairs.

4.6.3   Implementation

Calibration of Ancillary Moments

We compute cost and revenue parameters for monopoly newspapers with $Z_t \in [0.45, 0.55]$. We calibrate a_h to the average annual advertising revenue per copy and MC to the average annual variable cost per copy. Annual circulation revenue is typically below posted prices, partly because of discounts to subscribers. We compute the average discount as the average ratio of subscription price to annual circulation revenue, and apply this discount to all subscription prices to compute the effective

84

price of each newspaper. Appendix C.1 presents evidence on the sensitivity of our findings to variation in calibrated moments.

Pairing of Markets and Towns

Both our supply and demand models exploit spatial correlation in ideology to identify the unobservable component of ρ, the share of households that are Republican. This strategy requires that correlation in ρ be the only source of correlation in firms' and households' decisions across markets and towns that are paired together. On the supply side, this means pairing markets that are far enough apart that their newspapers do not compete directly. On the demand side, it means pairing towns that are far enough apart that the same exact newspapers are unlikely to be available in both towns in a pair.

To estimate the supply model, we require that paired markets be between 100 and 400 kilometers apart and located in the same state. Among possible market pairs, we identify the pair with lowest absolute difference in log population, breaking ties randomly. We then remove the matched pair from consideration and find the pair with the next closest population. We repeat this matching process until all pairs are matched.

Figure 4.1 illustrates the economic logic of our approach to pairing markets. Two counties located $100 - 400$ kilometers apart have a highly correlated Republican vote share and fraction white. However, due to physical transportation costs, newspapers headquartered in the first county rarely circulate in the second at such distances. Therefore, the correlation in firms' choices across markets located $100 - 400$ kilometers apart plausibly reflect the response to household characteristics, rather than a

direct competitive response to firms in neighboring markets.[3]

We use the same algorithm to pair towns for demand estimation that we use to pair markets for supply estimation. Here, the economic logic is similar: towns at such distances typically have non-overlapping sets of newspapers available. Therefore, at such distances, spatial correlation in households' demand for Republican and Democratic newspapers is likely to reflect unobservable heterogeneity in household ideology rather than, say, unmeasured variation in newspaper quality.

Computational Methods

We estimate via two-step maximum likelihood. We first estimate the demand model. We then estimate the supply model taking demand model parameters as given. We compute asymptotic standard errors using a numerical Hessian, adjusting for the use of a two-step procedure following Murphy and Topel (1985).

We approximate the likelihood via sparse grid integration with Gaussian kernel and accuracy 3 (Heiss and Winschel 2008, Skrainka and Judd 2011). In the online appendix, we present estimates of the model in which we reduce and increase the accuracy by 1.

We maximize the likelihood using KNITRO's active-set algorithm for unconstrained problems (Byrd, Nocedal, and Waltz, 2006). We use exponential transforms to ensure that all standard deviations are positive so that the likelihood is well-defined. In estimating the demand model, we use an exponential transform to constrain $\Gamma > 0$ (otherwise newspapers are complements). We also constrain parameters so that the predicted price and circulation share of a monopoly newspaper in a market with

---

[3] Common ownership of newspapers in different markets is another possible source of correlation. In appendix C.1 we show that removing the small number of market pairs with common ownership makes little difference to our results.

86
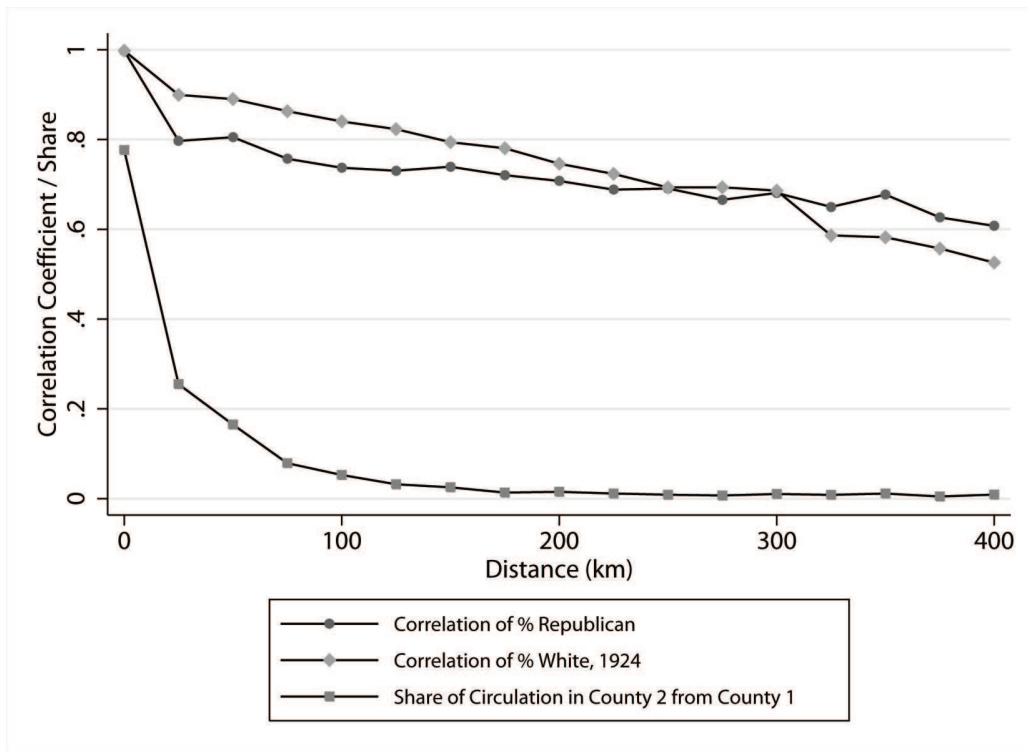
Figure 4.1: Spatial Decay in Newspaper Shipments and Demographic Correlations
Notes: The first two lines show the correlation coefficient of fraction Republican and fraction white for counties located in the same state, at different centroid distances. Republican vote share is the average Republican share of the two-party vote in presidential elections from 1868-1928. The third line shows the share of newspaper circulation in county 2 accounted for by newspapers headquartered in county 1, for counties located at different centroid distances. Only counties containing at least one sample market are included.

$\rho = 0.5$ is equal to the sample means for monopoly markets with $Z_t \in [0.45, 0.55]$.[4]

For demand estimation we choose starting values either at zero or at a value (typically one) reflecting the expected order of magnitude of the parameter. For supply estimation we begin with order-of-magnitude starts, and estimate two sub-models to improve the accuracy of the starting values supplied to the final estimator. The first sub-model is a post-entry version of the model that conditions on the number of newspapers entering each market. The second sub-model is an estimate of the entry game taking the post-entry parameters as given.

Evaluation of the supply model likelihood requires imposing equilibrium in the entry stage, affiliation choice stage, pricing stage, and advertising pricing stage. We provide above an explicit characterization of the equilibrium in the affiliation and advertising pricing stages. For given fixed costs κ and variable profit $V()$, the entry stage game admits a unique and explicit solution provided $V()$ is strictly decreasing in the number of entering newspapers. In repeated simulations we find that this property holds for all markets at the estimated parameters. The equilibrium of the pricing game is characterized by a system of first-order conditions, which we solve using MINPACK's (Mor, Garbow, and Hillstrom., 1980) implementation of the Powell (1970) hybrid method.[5] We choose a starting value close to the observed prices ($4) and verify that the solution is not sensitive to local variation (plus or minus $1 per copy) in the choice of starting value at the estimated parameters.

We set $\bar{J} = 3$ so that we treat affiliations as unobserved in markets with four or more newspapers. Only 8 markets in our data have four or more newspapers.

The online appendix presents Monte Carlo experiments and experiments with

---

[4] This constraint implies an explicit (closed form) solution for α and $\underline{\beta}$ as a function of the other parameters that is trivial to compute.

[5] We use the C/C++ implementation of MINPACK distributed by Frη|ˇdη|ˇric Devernay.

random starting values for both the demand and supply steps of the estimation.

4.7 Identification

In this section, we present a heuristic overview of the features of the data that identify the model's parameters. We begin with a heuristic discussion of the role of spatial correlation in identifying the incentive to differentiate. We then turn to a step-by-step discussion of the model stages.

4.7.1 Incentive to Differentiate

It is helpful to begin by considering the following reduced-form approximation of the model. Each market has two newspapers, which we refer to as the Incumbent and the Entrant. Newspapers successively choose affiliations in order of entry. A reduced-form profit function governs the payoff to each newspaper from choosing R relative to the payoff from choosing D.

The Entrant's payoff to choosing R is a function of household ideology, the Incumbent's affiliation, and an idiosyncratic shock. The Incumbent's payoff to choosing R is a function of household ideology and an idiosyncratic shock. (In the model we estimate, the Incumbent's payoff also incorporates the Incumbent's beliefs about the Entrant's choice of affiliation.)

The econometrician wishes to recover the extent to which the incentive to differentiate drives diversity. The econometrician observes newspapers' affiliations but not household ideology, which may vary across markets.

The incentive to differentiate depends on the Entrant's payoffs. If the Entrant's payoff to R is much greater when the Incumbent chooses D, then the incentive to differentiate will play an important role in determining equilibrium diversity. If the

Table 4.5: Affiliation Choices in Own and Neighboring Markets

Share of second entrants choosing Republican affiliation

| | Incumbent Affiliation | |
	Democratic	Republican
Incumbent Market:		
Own	.50	.53
Neighbor	.33	.66
Number of markets	269	

Notes: Data are from supply estimation sample and include all markets with at least two newspapers in which the neighboring market has at least one newspaper.

Entrant's payoff to *R* is independent of the Incumbent's choice, then diversity will not depend on competitive forces.

From equilibrium market configurations alone it will be difficult to recover the incentive to differentiate. Consider the data in the first row of table 4.5, which shows summary statistics on the affiliation choice of second entrants in our data. In markets where the Incumbent is *D*, the Entrant is *R* about half the time. In markets where the Incumbent is *R*, the Entrant is slightly *more* likely to be *R*.

Based on these data two conclusions are possible. The first is that the incentive to differentiate is weak. The second is that unmeasured variation in household ideology is driving both Incumbent and Entrant affiliations, leading to a slightly positive empirical correlation in affiliations that masks important competitive forces.

One solution to this problem is to condition on observable proxies for household ideology. As table 4.4 illustrates, that approach will lead to a significantly negative conditional correlation between Incumbent and Entrant affiliations. But, such an approach leaves open the possibility that the observable proxy does not capture all variation in household ideology. If it does not, estimates based on observed configurations will tend to understate the incentive to differentiate.

We will couple an observable measure of household ideology with an additional

source of information on the importance of unobservable variation in ideology: the spatial correlation in newspapers' affiliation choices. The second row of table 4.5 illustrates the logic of this approach. A given Entrant's choice of affiliation is strongly positively correlated with the choice of the Incumbent in a neighboring market. Because we construct pairs to minimize the chance of direct economic competition between neighbors, the natural interpretation of this correlation is that it reflects spatially correlated variation in household ideology.

If household ideology were unobserved but identical across neighboring markets, a fixed effects or differences-in-differences strategy would be sufficient to control for the confounding effect of ideology and recover the incentive to differentiate. Because an Entrant's affiliation choice is more positively correlated with its neighboring Incumbent's affiliation than with its own Incumbent's affiliation, such a fixed effects strategy would show a strong incentive to differentiate.

However, it is unlikely to be appropriate in general to assume that neighboring markets have identical household attributes. Such an assumption would be false for observed characteristics, which are highly, but imperfectly, correlated across neighbors. Instead of assuming perfect correlation of the unobservables, we assume the correlation in unobservables matches that of our observable proxy for ideology. Speaking loosely, this amounts to scaling up the correlation between the Entrant's affiliation and that of the neighboring Incumbent, and subtracting the scaled correlation from the correlation between the Entrant's affiliation and that of its own Incumbent.

4.7.2   *Supply Model*

Take the estimated demand system as given. We work backwards through the stages of the game.

Begin with the advertising stage. The parameter a_l governs the extent to which newspapers earn less on overlapping readers than singleton readers. Fixing other parameters, when $\bar{\beta}$ is large enough relative to $\underline{\beta}$, readership overlaps more between two newspapers that have the same affiliation than between two newspapers that have different affiliations. Therefore a_l, combined with the parameters of the demand system, determines the incentive to differentiate. Because the demand parameters are given, the parameter a_l can be thought of as identified by the extent to which newspapers differentiate more than would be expected from the demand system alone, i.e. more than would be expected if $a_l = a_h$ and hence newspapers did not compete on advertising.

The incentive to differentiate is, in turn, identified from the assumptions we make about the spatial correlation in the unobservables. These assumptions also identify σ_δ and σ_η, the parameters that govern the extent to which ideology varies across markets conditional on observables.

Move next to the pricing game. Here there are no parameters to estimate: given newspapers' affiliations, the pricing game is fully determined by the demand system. Note that, in this sense, the argument for identification of the advertising stage above is dependent on conduct assumptions for the pricing game.

Consider next the game in which newspapers sequentially choose affiliations. Expected payoffs come from the pricing and advertising stages. The extent of variation σ_ξ in cost shocks ξ are identified as an unexplained residual in newspapers' affiliation choices. The mean of the unobservable μ_δ is identified from the extent to which newspapers choose to be Republican "too often" given the parameters of the demand system and the observable fraction Republican in the market.

Move next to the entry game. Payoffs to entry as a function of the number of en-

trants are delivered by the stages above. These payoffs, in turn, identify the fixed cost cutoffs that determine the equilibrium number of entrants. The correlation between the number of newspapers and the market's population, and the extent of variation in the number of newspapers conditional on population, pin down the entry-stage parameters μ_κ^0, μ_κ^1 and σ_κ respectively.

Note that, because newspaper fixed costs are increasing in market size (Berry and Waldfogel, 2010), we cannot use the homogeneity assumption of Bresnahan and Reiss (1991) to identify the entry cutoffs directly. An important implication is that the identification of the entry stage partly "feeds back" into the identification of the later-stage parameters, which means that later-stage parameters are also influenced by the observed number of entrants and the fit of the entry model.

4.7.3 Demand Model

Suppose that there is no unobservable heterogeneity in town ideology, i.e. that $\sigma_\delta = \sigma_\eta = 0$ for towns. Then, fixing the affiliations of available newspapers, the correlation between the relative demand for Republican newspapers and the observed fraction Republican identifies $\overline{\beta}$ relative to $\underline{\beta}$. Given the relative magnitudes of these parameters, the share of households reading the newspaper in markets with known ideological composition pins down their absolute value. Given these two parameters, observed monopoly markups with known ideological composition identifies the price sensitivity parameter α.

Table 4.3 shows that, holding constant the observed fraction Republican, Republican newspapers on average get lower circulation in markets with more Republican newspapers available. That fact pins down the extent to which same-affiliation newspapers are substitutable in demand, which in turn identifies the remaining utility

parameter Γ. Given utility parameters, the parameter σ_ζ, which governs the importance of measurement error in circulation, is identified as the variance of residual circulation.

The relationship between the share of a town's available newspapers and the observed share Republican then identifies the parameters μ_ρ^0 and μ_ρ^1.

The preceding argument presumes that the econometrician perfectly observes the share of Republican households in each market. In practice there is likely to be some unmeasured heterogeneity in household ideology. Markets with more Republican households will tend to have more Republican newspapers available, which means that a naive estimator will tend to understate both the difference between $\overline{\beta}$ and $\underline{\beta}$ and the extent of substitution Γ.

We address this issue by exploiting the spatial correlation in circulation, in a manner similar to that outlined in section 4.7.1 above. To the extent that the relative circulation in a given town is positively correlated with the number of Republican newspapers available in a neighboring town (or with the circulation patterns in the neighboring town), we interpret that as evidence of correlated heterogeneity in household ideology. Spatial covariance patterns then identify σ_δ and σ_η, as in the supply model.

For this strategy to make sense, it is important that paired towns be far enough away that there is little direct economic interaction in their news markets. Otherwise, unmeasured correlation in, say, newspaper quality could lead us to overstate the importance of unobservables on the demand side.

4.8 Model Estimates

4.8.1 Model Estimates

Table 4.6 reports estimates of demand model parameters. The qualitative patterns are consistent with our economic intuition and with the descriptive evidence in table 4.3. Households prefer newspapers whose affiliations match their own. Same-type newspapers are substitutes in demand. There is unobservable heterogeneity in household ideology across towns, which in turn is correlated with the fraction of available newspapers that are Republican.

Table 4.7 reports estimates of supply model parameters. Consistent with our economic model we find that advertising rates are lower for overlapping readers than for singleton readers. We find some evidence of unobservable heterogeneity in household ideology, though it is less important than on the demand side.

Our model implies that the average newspaper receives $6 of circulation revenue and $11 of advertising revenue per reader per year (in 1924 dollars). Thus, consistent with contemporaneous evidence, advertising accounts for the majority of revenue. Variable costs are $8 per reader per year, and so variable profits are roughly $9 per reader per year. These profits are high, but a good share are dissipated in fixed costs such as editorial costs.

We estimate that the average newspaper sells 0.32 copies per household each day. Among households whose type is the majority in their market (R households in majority R markets or D households in majority D markets), this ratio rises to 0.35. For households whose political type is the minority in their market (D households in majority R markets or R households in majority D markets), the ratio falls to 0.27. Consistent with our reduced-form evidence, the match between a paper's affiliation

Table 4.6: Parameter Estimates: Demand Model

Price coefficient (α)	0.1802
	(0.0025)
Mean utility for different-affiliation paper ($\underline{\beta}$)	-0.1887
	(0.0592)
Mean utility for same-affiliation paper ($\bar{\beta}$)	0.7639
	(0.0664)
Substitutability between same-type papers (Γ)	0.2438
	(0.0562)
Standard deviation of log-measurement error (σ_q)	0.6995
	(0.0077)
Mean of unobservable shifter of fraction Republican (μ_δ)	0.0945
	(0.0545)
Standard deviation of unobservable ($\sqrt{\sigma_\delta^2 + \sigma_\eta^2}$)	0.2859
	(0.0133)
Parameters governing share of town's newspapers that are Republican	
μ_ρ^0	-0.1680
	(0.1098)
μ_ρ^1	2.0006
	(0.0338)

Calibrated parameters:	
Marginal cost (MC)	8.1749
Spatial correlation of unobservable ($R \equiv \frac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_\eta^2}$)	0.7286

Number of Unique Towns	12198
Number of Unique Newspapers	669
Number of Newspaper-Towns	28841

Notes: Table shows the estimated parameters of the demand model with asymptotic standard errors in parentheses.

Table 4.7: Parameter Estimates: Supply Model

Advertising revenue per reader of non-singleton bundles (a_l)	6.2100
	(0.6619)
Standard deviation of affiliation cost shocks (σ_{ξ})	0.2026
	(0.0170)
Mean of unobservable shifter of fraction Republican (μ_{δ})	-0.0183
	(0.0178)
Standard deviation of unobservable ($\sqrt{\sigma_{\delta}^2 + \sigma_{\eta}^2}$)	0.0956
	(0.0803)
Parameters governing the distribution of fixed costs	
μ_{κ}^0	8.5215
	(0.2860)
μ_{κ}^1	-0.6281
	(0.0391)
σ_{κ}	0.3519
	(0.0209)
Calibrated parameters:	
Advertising revenue per reader of singleton bundles (a_h)	13.2811
Spatial correlation of unobservable ($R \equiv \frac{\sigma_{\delta}^2}{\sigma_{\delta}^2 + \sigma_{\eta}^2}$)	0.7217
Number of Markets	1910
Number of Newspapers	1338

Notes: Table shows the estimated parameters of the supply model. The supply model is estimated taking the demand model parameters as given. Asymptotic standard errors in parentheses adjust for the two-step estimation procedure. The advertising rate a_h is calibrated as described in section 4.6.3.

and their consumers' ideology is an important determinant of newspaper demand.

In the online appendix, we present estimates of the main regression specifications in tables 4.3 and 4.4 using data simulated from the model at the estimated parameters. These regressions show that the estimated model fits key features of the data well.

4.8.2 *Determinants of Equilibrium Diversity*

Table 4.8 assesses how market forces determine the extent of political diversity in equilibrium. For our baseline model and a series of counter-factual models we perform 5 independent simulations of the affiliation choices of all newspapers in our empirical sample. We report the average across simulations of the share of multi-paper markets that are diverse. We define a newspaper market to be diverse if it has at least one Republican paper and one Democratic paper. At the estimated parameters, the model predicts that 58 percent of multi-paper markets are diverse.

In our first counterfactual, we assume that each entering newspaper chooses its affiliation as if it expected to be a monopolist in the market. The share of multi-paper markets that are diverse falls by nearly half, to 32 percent. The incentive to differentiate from competing papers is a powerful force encouraging diversity.

In our second counterfactual, we assume that each entering newspaper chooses its affiliation as if its market had equal numbers of R and D type households. The share of multi-paper markets that are diverse rises significantly, to 85 percent. The incentive to cater to households tastes significantly limits diversity.

In our third counterfactual, we assume that each entering firm chooses its affiliation as if $\xi = 0$. The cost shocks ξ are simply a residual in the model, but one can interpret them as capturing the preferences or fixed assets of owners, along with other idiosyncratic factors. Eliminating such factors would reduce the share of multi-paper

Table 4.8: Equilibrium Determinants of Diversity

	Share of multi-paper markets that are diverse
Baseline	0.59
When choosing affiliation, newspapers:	
Ignore competitors' choices	0.34
Ignore household ideology	0.83
Ignore idiosyncratic cost shocks (ξ)	0.42
Owners chosen at random from local households and newspaper type equals owner type	0.60

Notes: Table shows averages over 5 counterfactual simulations using the model estimates reported in tables 4.6 and 4.7. We define a market to have diverse papers if there is at least one Republican-affiliated paper and one Democrat-affiliated paper in this market. Counterfactuals are defined as follows. "Ignore competitors' choices" means that each entering newspaper chooses its affiliation as if it will be the only newspaper in the market. "Ignore household ideology" means that each entering newspaper chooses its affiliation as if exactly one-half of households are Republican ($\rho = 0.5$). "Ignore idiosyncratic cost shocks" means that each entering newspaper chooses its affiliation as if $\xi = 0$. "Owners chosen at random" means that a newspaper's affiliation is a random draw from the affiliations of households in its market. Number of newspapers is fixed at the value in the baseline simulation for all counterfactuals. Markets simulated to have five or more newspapers are treated as having five newspapers.

markets that are diverse from 58 percent to 42 percent: a nontrivial reduction, but not as large as the effect of ignoring competitors or of ignoring household preferences.

In our fourth and final counterfactual, we assume that newspaper owners are randomly chosen from the households in the market and a newspaper's affiliation is simply its owner's affiliation. Under this scenario, the share of multi-paper markets with diverse papers rises slightly from 58 percent to 60 percent. That is, economic forces result in diversity that is comparable to what would be observed if newspaper affiliations were chosen to be representative of households in the local market.

4.8.3 Model Specification and Implications for Diversity

Our model implies an important role for competition in generating ideological diversity in multi-paper markets. Table 4.9 illustrates the importance of allowing for heterogeneity in household ideology in reaching that conclusion. The table presents the ratio of the diversity share absent competition (if entering newspapers acted as monopolists) to the diversity share at baseline under four different modeling assumptions.

The first row presents estimates allowing for unobservables. The first column also includes our observable measure of the fraction Republican and is therefore equivalent to the specification reported in table 4.8. Diversity would decline by about half if newspapers acted as monopolists. The second column shows results from a specification in which we ignore the information contained in our observable measure of the fraction Republican. Strikingly, the estimated effect of competition on diversity is almost unchanged. This is especially noteworthy given the significant power of the observable fraction Republican to predict newspapers' affiliation choices, as illustrated in table 4.4.

Table 4.9: Model Specification and Implications for Diversity

Share of ideological diversity that would persist if newspapers ignored competitors' choices

		Include observable fraction Republican	
		Yes	No
Include unobservable	Yes	0.58	0.48
fraction Republican	No	0.60	0.97

Notes: Table shows results from simulations using various estimates of the supply model, taking as given the demand estimates from table 4.6. We define a market to have diverse papers if there is at least one Republican-affiliated paper and one Democrat-affiliated paper in this market. In each case we report the ratio of the fraction of diverse markets under the "ignore competitors' choices" counterfactual to the fraction at baseline. "Ignore competitors' choices" means that each entering newspaper chooses its affiliation as if it will be the only newspaper in the market. Including unobservables and observables yields the model estimates reported in table 4.7. "No unobservables" uses estimates from a constrained version of the model in which there is no unobservable heterogeneity in household ideology ($\sigma_\delta = \sigma_\eta = 0$). "No observables" uses estimates from a version of the model in which we assume that all markets have measured fraction Republican $Z_m = 0.5$. Markets simulated to have five or more newspapers are treated as having five newspapers.

Contrast these findings with those from the second row, where we assume that there is no unobservable variation in the fraction Republican (by setting $\sigma_\delta = \sigma_\eta = 0$). When we include the information contained in observables, the result is similar to our main specification. But when we ignore the information contained in observables, the model returns the answer that competition plays no role in fostering diversity. The finding is intuitive: as in table 4.4, absent controls for household ideology, there is only a weak empirical correlation between an entering newspaper's affiliation and that of its incumbents. The model interprets this to mean that advertising competition is weak (a_l is near to a_h) and hence that newspapers have only a limited incentive to differentiate on ideology.

In a model that assumes no unobservable cross-market heterogeneity in household ideology, counterfactual implications for diversity are highly dependent on the

researcher's access to appropriate observable proxies for market ideology. By contrast, exploiting spatial correlation to allow for unobservable heterogeneity in household ideology results in a model that is far more robust to variation in the quality of observable ideology measures.

4.9 Policy Simulations

4.9.1 Definitions

We evaluate two types of government policies in our counterfactual simulations.

The first type is a relaxation of antitrust rules along the lines of joint operating agreements. Joint operating agreements have existed since at least 1933 and were given formal exemption from antitrust enforcement action in the Newspaper Preservation Act of 1970 (Busterna and Picard, 1993). The Act states its goal as "maintaining a newspaper press editorially and reportorially independent and competitive in all parts of the United States." The act allows approved newspapers, in essence, to collude on prices and advertising rates provided that they remain editorially independent. Joint operating agreements have subsequently been approved selectively in some US cities; in our simulations we assume they are operative everywhere.

We define a joint operating agreement as an arrangement with both price and advertising collusion. We define collusion to consist of setting prices or advertising rates to maximize the sum of profits of all entering newspapers. Under collusion, we assume that each newspaper chooses its affiliation independently without regard to the profits of other newspapers. Each newspaper expects that it will keep all of its subscription revenue and that it will share advertising revenue in proportion to its circulation.

Formally, we define a collusive price of newspaper j as the jth element of a price

vector p^* that solves

$$p^* \in \operatorname*{argmax}_p \sum_j \left(p_j + a_j \left(p \right) - MC \right) q_j \left(p \right) \tag{4.23}$$

where here we make explicit the dependence of advertising rates and demand on the full vector of prices. We define the collusive per-reader advertising revenue of newspaper j as

$$a_j = a_h \left(\frac{1 - q_0}{\sum_k q_k} \right) + a_l \left(1 - \frac{1 - q_0}{\sum_k q_k} \right) \tag{4.24}$$

where q_0 is the share of households that read no newspaper.[6]

These assumptions are a reasonable match to the revenue-sharing arrangements of joint operating agreements authorized under the Newspaper Preservation Act (Busterna and Picard, 1993). In some cases a newspaper's share of revenue is a "sliding" function of the newspaper's contribution to revenue or to total advertising sales. In other cases, the revenue sharing rule is fixed in advance, but in such cases is usually related to the initial capital investment of the newspapers, and hence to their financial health at the time of the agreement. In both types of arrangements, a newspaper with a greater circulation will generally be entitled to a greater share of the joint venture's revenue.

The second type of government policy that we evaluate is newspaper subsidies. We first consider the impact of eliminating postal subsidies to newspapers. In 1924, the post office's cost of publication delivery exceeded its revenue by a factor of more than three (Kielbowicz, 1994). Assuming these subsidies apply equally to all postal deliveries, we estimate that the marginal cost of the average newspaper would have

---

[6] The per-household value of advertising across all newspapers is given by $a_h \left(1 - q_0 \right) + \left(\sum_k q_k - \left(1 - q_0 \right) \right) a_l$.

risen by 15 percent if postage were charged at cost. We therefore define elimination of postal subsidies to mean increasing *MC* to 1.15 times its calibrated value.

We next consider a subsidy modeled after the system of newspaper subsidies in Sweden, which favors a local market's "second papers," i.e. papers with lower circulation than the largest paper in the market. We implement the subsidy as a fixed payment to all second entrants. Formally, we assume that a second entrant earns $V(2) + K$ where K is the amount of the subsidy. We set K equal to 15 percent of pre-subsidy revenue to match the approximate share of second-paper revenue coming from subsidies in Sweden (Gustaffson, rnebring, and Levy, 2009).

4.9.2 Results

Tables 4.10 and 4.11 present simulations of the effect of various government policies. We report the effect of these policies on market structure and diversity in table 4.10 and on welfare in table 4.11.

As table 4.10 shows, joint operating agreements increase equilibrium diversity. This is the result of two countervailing effects. Conditional on the number of firms in the market, joint operating agreements soften the competitive incentive to differentiate and thus make diverse configurations less likely. Thus, the share of two-firm markets with diverse papers falls from 42 percent to 37 percent and the share of markets with three or more firms that are diverse falls from 79 percent to 76 percent. At the same time, joint operating agreements encourage entry and thus increase the number of markets with multiple firms. This was the primary motivation for the Newspaper Preservation Act, and we find this effect is large: the number of markets with two firms increases from 146 to 212 and the number of markets with three or more firms increases from 108 to 258. On net, the effect of increased entry on diver-

Table 4.10: Government Policy and Diversity

Markets with:	0 Firm (#)	1 Firm (#)	2 Firms (#)	2 Firms (% Diverse)	3+ Firms (#)	3+ Firms (% Diverse)	Newspaper Markets (% Diverse)	Share of Hhlds with Diverse Papers
Baseline	951	709	143	0.43	106	0.80	0.15	0.23
Collusion								
Pricing only	951	682	143	0.38	134	0.77	0.16	0.25
Advertising only	951	499	209	0.41	250	0.78	0.29	0.36
JOAs	951	492	200	0.40	266	0.77	0.30	0.37
Subsidies								
No postal subsidy	1387	387	80	0.44	56	0.80	0.15	0.15
2nd entrant subsidy	951	431	422	0.43	106	0.80	0.28	0.32

Notes: Table shows averages over 5 counterfactual simulations using the model estimates reported in tables 4.6 and 4.7. We define a market to have diverse papers if there is at least one Republican-affiliated paper and one Democrat-affiliated paper in this market. Collusion is defined as setting prices or ad rates to maximize the joint (total) profits of all entering newspapers. Markets simulated to have five or more newspapers are treated as having five newspapers.

105

Table 4.11: Government Policy and Welfare

	Avg. Price in Multi-Firm Mkts	Avg. Ad Revenue in Multi-Firm Mkts	Per-Household Surplus			
			Consumers	Firms	Advertisers	Total
		(All values are annual totals in 1924 dollars)				
Baseline	6.24	10.34	3.28	0.42	0.50	4.20
Collusion						
Pricing only	7.93	10.72	2.90	0.44	0.39	3.73
Advertising only	5.82	11.43	4.60	0.60	0.00	5.20
Joint operating agreements	6.80	11.57	4.06	0.65	0.00	4.71
Subsidies						
Eliminate postal subsidy	6.94	10.58	2.04	0.25	0.26	2.54
Subsidy for second entrant	6.00	10.74	3.63	0.21	0.63	4.47

Notes: Table shows averages over 5 counterfactual simulations using the model estimates reported in tables 4.6 and 4.7. We define a market to have diverse papers if there is at least one Republican-affiliated paper and one Democrat-affiliated paper in this market. Collusion is defined as setting prices or ad rates to maximize the joint (total) profits of all entering newspapers. Markets simulated to have five or more newspapers are treated as having five newspapers.

sity dominates the effect of decreased differentiation. The share of all markets with at least one newspaper that have diverse papers increases from 15 percent to 28 percent and the share of households living in a market with diverse papers increases from 22 percent to 35 percent.

The table shows that subsidies also increase diversity. Eliminating postal subsidies and adding subsidies for second entrants have small effects on differentiation conditional on market structure but large effects on entry. On net, the share of households in markets with diverse papers falls to 13 percent in the former case, and increases to 28 percent in the latter case.

Table 4.11 shows that joint operating agreements have a net positive effect on both consumer and producer surplus. Average consumer surplus per household rises from $3.25 to $4.45 and average firm profit per household rises from $0.28 to $0.41. Households benefit from collusion in part because of greater entry, and in part because of the two-sided nature of the market: higher advertising prices encourage newspapers to lower their prices to attract more readers. Total surplus increases from $4.02 to $4.85 per household.

Our findings regarding the effect of joint operating agreements are consistent with the limited empirical literature on the subject. Busterna and Picard (1993) conclude that there is little evidence of significant effects on consumer prices or on newspaper content, but at least some evidence that joint operating agreements lead to higher advertising rates (or faster growth in advertising rates) than would prevail in competitive markets.

Table 4.11 shows that subsidies increase consumer surplus. Firm surplus falls due to increased entry. Advertiser surplus rises with increased entry since advertisers only earn positive surplus in multi-paper markets. Note that in the case of the second-

entrant subsidy the firm surplus numbers do not include the value of the subsidy itself, and in neither subsidy case do the total surplus numbers reflect the cost of the subsidy to the government.

4.10   Conclusions

We estimate a model of newspaper partisanship in which partisanship affects household demand and is treated as a strategic decision by entering newspapers. We find evidence that partisanship influences the composition of readership and that it affects patterns of substitution among competing papers. We find, in turn, that entering newspapers take competitors' partisan affiliations into account when choosing their own. The model implies that competition is a crucial determinant of ideological diversity in media markets, and permits simulation of a number of counterfactual experiments that are relevant to contemporary policy debates.

APPENDIX

A. APPENDIX TO CHAPTER 2

A.1 Derivation of Hotelling Case

In the Hotelling case, consumer utility from the final good takes the form

$$
\begin{aligned}
u_{Ai} &= \delta_A - p_A - \theta_i \\
u_{Bi} &= \delta_B - p_B - (1 - \theta_i)
\end{aligned}
$$

Demand for each good at prices p_A, p_B is given by integrating over the uniform distribution of types,

$$
\begin{aligned}
D_A\left(p_A, p_B\right) &= Pr\left(\delta_A - p_A - \theta_i > \delta_B - p_B - (1 - \theta_i)\right) \\
&= Pr\left(\theta_i < \frac{\delta_A - \delta_B + p_B - p_A + 1}{2}\right) \\
&= \frac{\delta_A - \delta_B + p_B - p_A + 1}{2} \\
D_B\left(p_A, p_B\right) &= \frac{\delta_B - \delta_A + p_A - p_B + 1}{2}
\end{aligned}
$$

Throughout we will assume that the equilibrium lies in the interior. This is satisfied whenever

$$
1 + p_A - p_B > \delta_A - \delta_B \ > p_A \ -p_B - 1
$$

In the common agency case, downstream firms charge no markups and so upstream firms set the wholesale prices to be the profit-maximizing retail prices:

$$\pi_A^C = (q_A - c) D_A (p_A = q_A, p_B = q_B)$$
$$\pi_B^C = (q_B - c) D_B (p_A = q_A, p_B = q_B)$$

First-order conditions for profit maximization are given by

$$q_A = \frac{\delta_A - \delta_B + q_B + 1 + c}{2}$$
$$q_B = \frac{\delta_B - \delta_A + q_A + 1 + c}{2}$$

The equilibrium is therefore given by wholesale and retail prices of

$$q_A^{C*} = p_A^{C*} = \frac{1}{3} (\delta_A - \delta_B) + 1 + c$$
$$q_B^{C*} = p_B^{C*} = \frac{1}{3} (\delta_B - \delta_A) + 1 + c$$

Profits to the upstream firms in equilibrium are thus

$$\pi_A^{C*} = \frac{1}{18} (\delta_A - \delta_B + 3)^2$$
$$\pi_B^{C*} = \frac{1}{18} (\delta_B - \delta_A + 3)^2$$

In the exclusive case, the exclusive carrier chooses a price to maximize profits given the wholesale price q_A:

$$\pi_w^E = (p_A - q_A) D^A (p_A, p_B = q_B)$$
$$p_A = \left(\frac{1 + \delta_A - \delta_B + p_B + q_A}{2} \right)$$

To avoid double marginalization, Firm A will offer a two-part tariff with wholesale price equal to marginal cost and a tariff equal to all of the profits. The two upstream firms profits are given by:

$$
\begin{aligned}
\pi_A^E &= \left(\frac{1 + \delta_A - \delta_B + p_B + c}{2} - c \right) D^A \left(p_A = \left(\frac{1 + \delta_A - \delta_B + p_B + c}{2} \right), p_B = q_B \right) \\
\pi_B^E &= (q_B - c) D^B \left(p_A = \left(\frac{1 + \delta_A - \delta_B + p_B + q_A}{2} \right), p_B = q_B \right)
\end{aligned}
$$

Firm B's optimal wholesale price rises now, leading to a higher retail price as well:

$$
q_B^{E*} = p_B^{E*} = c + \frac{3}{2} + \frac{1}{2} \left(\delta_B - \delta_A \right)
$$

Equilibrium profits when A is exclusive and B is not are given by

$$
\begin{aligned}
\pi_A^{E*} &= \frac{1}{32} \left(\delta_A - \delta_B + 5 \right)^2 \\
\pi_B^{E*} &= \frac{1}{16} \left(\delta_B - \delta_A + 3 \right)^2
\end{aligned}
$$

Finally, consider the case when Firm B is also exclusive, which we will denote by *EE*. Now two carriers set final retail prices to maximize their profits according to

$$
\begin{aligned}
\pi_{wA}^E &= (p_A - q_A) D^A (p_A, p_B) \\
\pi_{wB}^E &= (p_B - q_B) D^B (p_A, p_B)
\end{aligned}
$$

Solving, the equilibrium prices they will set as a function of wholesale prices are

$$
\begin{aligned}
p_A^{EE*} &= \frac{\delta_A - \delta_B + 2q_A + q_B}{3} + 1 \\
p_B^{EE*} &= \frac{\delta_B - \delta_A + 2q_B + q_A}{3} + 1
\end{aligned}
$$

Similar to above, we have that both A and B set two-part tariffs to avoid marginalization, and so set wholesale prices to marginal cost and earn tariff profits of

$$\pi_A^{EE} = \left(\frac{\delta_A - \delta_B + 2q_A + q_B}{3} + 1 - c \right) D^A \left(\frac{\delta_A - \delta_B + 2q_A + q_B}{3} + 1, \frac{\delta_B - \delta_A + 2q_B + q_A}{3} + 1 \right)$$

$$\pi_B^{EE} = \left(\frac{\delta_B - \delta_A + 2q_B + q_A}{3} + 1 - c \right) D^B \left(\frac{\delta_A - \delta_B + 2q_A + q_B}{3} + 1, \frac{\delta_B - \delta_A + 2q_B + q_A}{3} + 1 \right)$$

Optimizing, the two firms maximize profits, resulting in the following equilibrium:

$$q_A^{EE*} = c + 1 + \frac{1}{5} (\delta_A - \delta_B)$$

$$p_A^{EE*} = c + 2 + \frac{2}{5} (\delta_A - \delta_B)$$

$$\pi_A^{EE*} = \frac{1}{25} (\delta_A - \delta_B + 5)^2$$

Firm B's outcome is symmetric to this (swapping δ_A and δ_B).

A.2   Proofs for General Case

The following assumptions stand throughout:

1. Tastes for handsets are independent of tastes for carriers.

2. Handsets A and B are substitutes and their prices are strategic complements.

3. The upstream firms set wholesale prices and tariffs independently (i.e. no collusion is possible).

4. Share functions are continuous and differentiable in all prices. Pricing equilibria exist and are unique.

5. For simplicity, I will assume that the underlying demand system captures downstream "market power" with a parameter $\eta \in [0, \infty)$, such that under common agency, when $\eta = 0$, downstream firms are homogenous as in the above section so that for carrier n, $\frac{\partial s_{An}}{\partial p_{An}} = -\infty$. As η increases, so does $\frac{\partial s_{An}}{\partial p_{An}}$, and in the limit $\frac{\partial s_{An}}{\partial p_{An}} \to \frac{\partial s_A}{\partial p_A}$ as $\eta \to \infty$. This allows us to characterize the limit cases of carrier monopolists ($\eta = \infty$), carriers as homogenous ($\eta = 0$), and cases in-between. The analogous values for cross-partials are that $\frac{\partial s_{An}}{\partial p_{An'}}$ goes from ∞ to 0 as η goes from zero to ∞.

An example of a demand system that would satisfy A5: if consumers have taste draws θ_j for each firm $j = 1..J$, drawn from distributions F_j, and utility from the downstream good of firm j were of the form $u_{ij} = \kappa + \eta\theta_j - p_j$ for some constant κ. This is, in effect, a more general version of a Hotelling model. Note that a demand system of the Logit family would not satisfy this assumption, as downstream firms are always imperfect substitutes in that setting, and so the limit cases are not attainable.

One challenge is that as downstream firms gain more market power, total market power and the equilibrium prices increase, making direct comparisons of equilibrium prices for different levels of downstream market power difficult. For example, when carriers are monopolists, we would expect the carriers to retain some of the joint surplus; it would be unreasonable to expect that handset firms could extract the complete amount of joint surplus. Therefore, to simplify the comparisons, we will assume that when bargaining over the joint surplus, the outside alternative is to have the upstream firms sell handsets directly to consumers. This allows us to characterize the maximum surplus achievable by the upstream firms as the "direct" profits whenever joint profits are greater than that.

We will first analyze the common-agency case, where each carrier $n = 1..N$ offers both handsets. We will look for a symmetric equilibrium outcome. The upstream firms choose the wholesale prices q_{An} and q_{Bn} (and can further extract surplus from a flat tariff). Downstream firms choose final retail prices p_{An} and p_{Bn}, $n \in \{1, ..., N\}$ according to

$$\pi_n = (p_{An} - q_{An}) s_{An} (p_{An}, p_{-An}) + (p_{Bn} - q_{Bn}) s_{Bn} (p_{Bn}, p_{-Bn}) \qquad \text{(A.1)}$$

Maximizing downstream profits yields two first-order conditions that must be satisfied for both carriers at the optimal retail prices p_A^{C*}, p_B^{C*}:

$$(p_{An} - q_{An}) = \left(-\frac{\partial s_{An} (\mathbf{p})}{\partial p_{An}} \right)^{-1} \left(s_{An} (p_{An}, p_{-An}) + (p_{Bn} - q_{Bn}) \frac{\partial s_{Bn} (\mathbf{p})}{\partial p_{An}} \right)$$

$$(p_{Bn} - q_{Bn}) = \left(-\frac{\partial s_{Bn} (\mathbf{p})}{\partial p_{Bn}} \right)^{-1} \left(s_{Bn} (p_{Bn}, p_{-Bn}) + (p_{An} - q_{An}) \frac{\partial s_{An} (\mathbf{p})}{\partial p_{Bn}} \right)$$

Notice that the share derivatives must take into account the indirect effect of prices on competing prices, since we have assumed that prices are strategic complements. For example, we have

$$\frac{\partial s_{An} (\mathbf{p})}{\partial p_{An}} = \frac{\partial s_{An}}{\partial p_{An}} + \frac{\partial s_{An}}{\partial p_{Bn}} \frac{\partial p_{Bn}}{\partial p_{An}} + (N - 1) \left(\frac{\partial s_{An}}{\partial p_{An'}} \frac{\partial p_{An'}}{\partial p_{An}} + \frac{\partial s_{An}}{\partial p_{Bn'}} \frac{\partial p_{Bn'}}{\partial p_{An}} \right) \text{ (A.2)}$$

$$\frac{\partial s_{Bn} (\mathbf{p})}{\partial p_{An}} = \frac{\partial s_{Bn}}{\partial p_{Bn}} \frac{\partial p_{Bn}}{\partial p_{An}} + \frac{\partial s_{Bn}}{\partial p_{An}} + (N - 1) \left(\frac{\partial s_{Bn}}{\partial p_{An'}} \frac{\partial p_{An'}}{\partial p_{An}} + \frac{\partial s_{Bn}}{\partial p_{Bn'}} \frac{\partial p_{Bn'}}{\partial p_{An}} \right) \text{ (A.3)}$$

where we make use of the fact that we are looking for symmetric equilibria to simplify. Since prices are strategic complements, all derivatives of prices with respect

to other prices are positive. We can immediately analyze the limit cases of down-stream competition: if carrier demand is perfectly elastic ($\eta = 0$), cross-carrier partial derivatives are infinite, resulting in zero markups. The resulting market outcome is identical to that where the upstream firms compete directly for consumers: handset makers effectively set the final price since q_A and q_B are passed through directly to consumers as p_A and p_B, resulting in equilibrium handset markups under common agency given by

$$\left(q_A^{C*} - c\right) = \left(-\frac{\partial s_A}{\partial p_A}\right)^{-1} s_A\left(\mathbf{p}^{C*}\right)\bigg|\, p_A = q_A, p_B = q_B$$

$$\left(q_B^{C*} - c\right) = \left(-\frac{\partial s_B}{\partial p_B}\right)^{-1} s_B\left(\mathbf{p}^{C*}\right)\bigg|\, p_A = q_A, p_B = q_B$$

Profits for the upstream firms are then

$$\pi_A^{C*} = \left(-\frac{\partial s_A}{\partial p_A}\right)^{-1} N s_{An}\left(\mathbf{p}^{C*}\right)^2 = \pi_B^{C*}$$

In the other limit case where downstream firms are monopolists (and so each carrier effectively serves a different "market"), we have $\eta = \infty$ and zero cross-carrier effects, and are left with only the first two terms of equations A.2 and A.3. The carrier then maximizes the joint profits as though the upstream firms were colluding (the carrier effectively vertically integrates with both upstream firms); these profits are maximized when handset manufacturers offer marginal cost pricing to eliminate the double-marginalization ($q_A = q_B = c$) and instead extract surplus through a tariff. Total profits are greater than in the previous limit case, although the upstream firms would not be able to extract the full surplus without actually colluding in setting wholesale prices, which we assume is not possible. Following the bargaining

116

assumption made above, the monopolist carrier retains at least the surplus created from internalizing both upstream firms' profits, the upstream firms are left with maximal profits of π_A^{C*} and π_B^{C*}.

In the intermediate cases, we can assume that upstream firms are effectively able to choose the final retail price as they know the markup function used by carriers and are free to set any wholesale price. The combination of variable profits and tariffs can not exceed π_A^{C*} due to the bargaining assumption (i.e. carriers retain surplus generated by their market power).

Now consider the case of exclusivity: handsets A and B are exclusive to carriers 1 and 2, respectively. The equilibrium first-order conditions for optimal prices p_A^{EE*} and p_B^{EE*} are now

$$(p_{A1} - q_{A1}) = \left(-\frac{\partial s_{A1}}{\partial p_{A1}}\right)^{-1} (s_{A1}(p_{A1}, p_{B2}))$$

$$(p_{B2} - q_{B2}) = \left(-\frac{\partial s_{B2}}{\partial p_{B2}}\right)^{-1} (s_{B2}(p_{A1}, p_{B2}))$$

As η goes from zero to ∞, we have that $\frac{\partial s_{A1}}{\partial p_{A1}}$ goes from $\frac{\partial s_A}{\partial p_A}$ to $\frac{\partial s_1}{\partial p_1}$. The handset competition dominates at low η, and the carrier competition dominates at high η.

Define these markup functions as $m(q_{A1}, q_{B2})$ and note that the markup is decreasing in own wholesale price but increasing in opposite wholesale price. Upstream firms, anticipating this markup function, now choose wholesale prices to maximize joint profits, according to

$$\pi_A^{EE} = (q_{A1} + m_{A1}(q_{A1}, q_{B2}) - c)\, s_{A1}(q_{A1} + m_{A1}(q_{A1}, q_{B2}), q_{B2} + m_{B2}(q_{A1}, q_{B2}))$$

$$\pi_B^{EE} = (q_{B2} + m_{B2}(q_{A1}, q_{B2}) - c)\, s_{B2}(q_{A1} + m_{A1}(q_{A1}, q_{B2}), q_{B2} + m_{B2}(q_{A1}, q_{B2}))$$

Optimizing, we get Firm A's first-order condition given by

$$q_A - c = -m_A + \frac{\left(1 + \frac{\partial m_A}{\partial q_A}\right) s_{A1}}{-\left(\frac{\partial s_{A1}}{\partial p_{A1}} \left(1 + \frac{\partial m_A}{\partial q_A}\right) + \frac{\partial s_{A1}}{\partial p_{B2}} \frac{\partial m_B}{\partial q_A}\right)}$$

Note that this simplifies to the the first-order condition from the homogenous carrier case if prices are not strategic complements (if there is no positive effect from $\frac{\partial m_B}{\partial q_A}$). Therefore, in the limit case of $\eta = 0$, equilibrium prices are higher when prices are strategic complements. Finally, profits for Firm A in this case are

$$\pi_A^{EE*} = \left(\frac{\left(1 + \frac{\partial m_A}{\partial q_A}\right)}{-\left(\frac{\partial s_{A1}}{\partial p_{A1}} \left(1 + \frac{\partial m_A}{\partial q_A}\right) + \frac{\partial s_{A1}}{\partial p_{B2}} \frac{\partial m_B}{\partial q_A}\right)}\right) s_{A1} \left(p_{A1}^{EE*}, p_{B2}^{EE*}\right)^2$$

Exclusivity is optimal iff

$$\overset{\pi_A^{EE*}}{\left(\frac{\left(1 + \frac{\partial m_A}{\partial q_A}\right)}{-\left(\frac{\partial s_{A1}}{\partial p_{A1}} \left(1 + \frac{\partial m_A}{\partial q_A}\right) + \frac{\partial s_{A1}}{\partial p_{B2}} \frac{\partial m_B}{\partial q_A}\right)}\right) s_{A1} \left(\mathbf{p}^{EE*}\right)^2} \overset{>}{-} \left(-\frac{\partial s_A}{\partial p_A}\right)^{-1} N s_{An} \left(\mathbf{p}^{C*}\right)^2 > \overset{\pi_A^{C*}}{0} \quad \text{(0A.4)}$$

We know that

$$\left(\frac{\left(1 + \frac{\partial m_A}{\partial q_A}\right)}{-\left(\frac{\partial s_{A1}}{\partial p_{A1}} \left(1 + \frac{\partial m_A}{\partial q_A}\right) + \frac{\partial s_{A1}}{\partial p_{B2}} \frac{\partial m_B}{\partial q_A}\right)}\right) > \left(-\frac{\partial s_A}{\partial p_A}\right)^{-1}$$

holds for all finite η, and that they are equal in the limit as $\eta \to \infty$ (there is no strategic complementarity of prices "across markets", or $\frac{\partial m_B}{\partial q_A} = 0$ in that limit). Also, for any given price vector $\mathbf{p}$, we have that $s_{A1}(\mathbf{p}) = N s_{An}(\mathbf{p})$ when $\eta = 0$, but $N s_{An}(\mathbf{p}) - s_{A1}(\mathbf{p})$ increases as η increases. That is, the amount of foregone sales from

exclusivity increases as consumers are less willing to substitute between downstream goods. We also know that equation 4 holds at $\eta = 0$. Combining these, we have that equation 4 holds at $\eta = 0$, but that the LHS is decreasing as η increases, and that equation 4 does not hold in the limit as $\eta \to \infty$. Under the continuity assumption, we can apply the intermediate value theorem to get that there exists an η^* at which point equation 4 holds with equality. Therefore, for all values of $\eta < \eta^*$, exclusivity is the profit maximizing strategy.

To address Proposition 2, we start with a model of what a carrier's willingness to pay is. For carrier $n \in \{1,2\}$, the alternative to having handset A exclusively is that carrier n' will have handset A exclusively (I will assume there is a handset B available to both carriers). The equilibrium outcome will be the one that maximizes the joint profits of the exclusive carrier and Firm A.

I first make a simplifying assumption: each carrier chooses only a network access price; handset prices are fixed across carriers at p_h. This simplifies the analysis, and I do not believe this to be a controversial assumption, as in November 2011 when the iPhone is available on three carriers, the device is priced identically across carriers but monthly access fees differ. The two carriers will have identical marginal costs c, and choose their monthly access prices p_n, which creates a final good price for handset h on carrier n of $p_n + p_h$. Carriers choose their monthly access price in the standard profit maximization framework. From now on, p_1 and p_2 represent equilibrium monthly access prices less marginal cost.

Each carrier's willingness to pay is determined by the difference in profits from having exclusivity versus its rival having exclusivity. I denote carrier 1 having exclusivity of handset A by $\chi = 1$, and carrier 2 having exclusivity with $\chi = 2$. For carrier 1, the willingness to pay to Firm A is therefore

$$p_1 (\chi = 1) \cdot (s_{A1} (\chi = 1) + s_{B1} (\chi = 1)) - (p_1 (\chi = 2) + p_A) \cdot (s_{B1} (\chi = 2))$$

Similarly, for carrier 2, it is

$$p_2 (\chi = 2) \cdot (s_{A2} (\chi = 2) + s_{B2} (\chi = 2)) - (p_2 (\chi = 1) + p_A) \cdot (s_{B2} (\chi = 1))$$

Re-arranging, we have each carrier's willingness to pay having two components: a change in profits from B, and the sales potential of A.

$$[p_1 (\chi = 1) \cdot s_{B1} (\chi = 1) - p_1 (\chi = 2) \cdot s_{B1} (\chi = 2)] + (p_1 (\chi = 1) + p_A) \cdot s_{A1} (\chi = 1)$$

$$[p_2 (\chi = 2) \cdot s_{B2} (\chi = 2) - p_2 (\chi = 1) \cdot s_{B2} (\chi = 1)] + (p_2 (\chi = 2) + p_A) \cdot s_{A2} (\chi = 2)$$

We are assuming that carrier 1 faces more elastic demand from its network. Therefore, at $\beta = 0$, we know that the first term for carrier 1 is larger than for carrier 2, and the difference is increasing in β. Further, we know that the second component is larger for carrier 2, since he has a higher quality network, and that this difference is growing in β. Therefore, to establish Proposition 2, we need to show that the 2nd component grows faster in β. This follows form the inclusion of p_A, which is fixed for all β. The price p_A is perfectly inelastic, whereas the equilibrium network prices cannot be, and so there reaches a point at which the limited market achievable by

carrier 1 dominates the gains carrier 1 can earn in monthly fees.

B. APPENDIX TO CHAPTER 3

B.1   Summary Statistics

Table B.1 shows summary statistics from the demand dataset.

B.2   Reduced-Form Evidence

First, Table B.2 shows a regression to show that consumers do indeed respond to network quality differences.

Figure B.1 shows raw shares across markets for carriers and smartphones.

Figure B.2 shows residuals from regressions of the market-level shares of carriers and smartphones on a set of controls, including network quality and income distributions.

B.3   Alternative Logit Approach

The model described in Section 4 is similar to the Pure Characteristics model described by Berry & Pakes (2007), which omits i.i.d. Logit draws for each possible good and opts instead for only random coefficients to rationalize tastes. A Logit approach in this setting would consist of adding an i.i.d. Logit errors to each discounted flow utility U_{imnht} and directly estimating a likelihood for each survey respondent. For example, if we observe a survey respondent that owns an iPhone on AT&T which was

Table B.1: Summary Statistics

	Main Sample
Number of Markets	90
Number of Months	26
Total Observations	573,121
Monthly Respondents: Minimum	18,836
Monthly Respondents: Maximum	24,030
Average monthly share who own no mobile phone	7.50%
Average monthly rate of smartphone purchase	1.36%

	Main Sample (Weighted)	Census
% Female	51.97%	52.06%
% of Adult Population Age 60+	25.54%	24.37%
% Income $100K+	17.22%	15.73%

*Table B.2:* Effect of Dropped Calls on Market Share
Dependent Variable: Market Share

	Specification 1		Specification 2	
Parameter	Estimate	S.E.	Estimate	S.E.
Dropped Calls	−0.8393	0.2089	−0.9924	0.3573
Carrier 1	0.05204	0.01096	0.2129	0.00873
Carrier 2	0.1668	0.01460	0.3572	0.0142
Carrier 3	0.1398	0.01491	0.3209	0.0128
Carrier 4	0.003632	0.01012	0.1537	0.00865
Constant	0.1230	0.009399	-	-
N	419		362	
R^2	0.4453		0.8741	

Results are from an OLS regression. Specification (1) includes all other carriers and the outside option in the "constant", whereas Specification (2) uses each carrier's share of the market held by the four national carriers (i.e. each national carrier's "inside share"). Standard errors are clustered at the market level. The data are for the 6th month of survey data. The constant represents consumers on minor carriers or without phones.
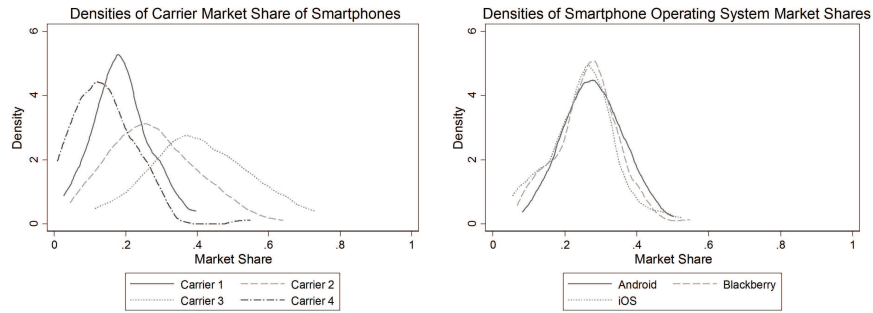
Figure B.1: Across-Market Variance in Shares of Carriers vs Smartphones
Note: shares are averaged over final three months of sample to reduce sample noise in smaller markets.
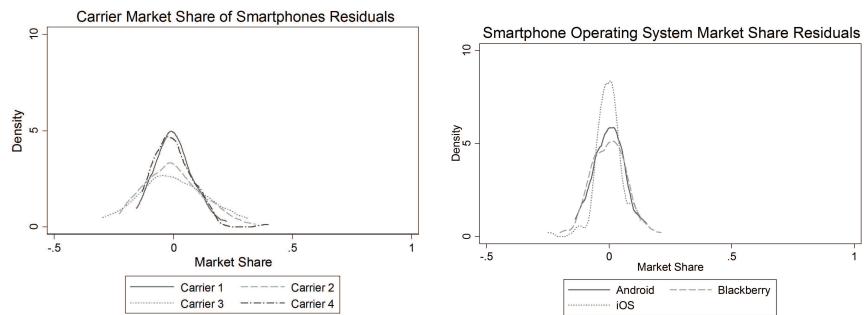
Figure B.2: Across-Market Residuals from Controlled Regressions
Note: shares are averaged over final three months of sample to reduce sample noise in smaller markets. Controls include income distributions and network quality (for carriers) and AT&T market share (for smartphones).

purchased 5 months ago, then we know that in the survey month, this consumer's state was a 4-month old iPhone on AT&T with 20 months remaining on contract and an early termination fee of, say, $155. We also know that in the survey month, this respondent chose to stay with their iPhone instead of switching to another device or network. We could model the Logit probability of this choice, and maximize the sum of the log likelihoods of these probabilities for all observations. Such an approach has multiple challenges in implementation:

First, such a setup would not easily allow for unobserved tastes (such as random coefficients) beyond the Logit draw. The reason for this is that unobserved taste vectors would have to be drawn from the conditional distribution based on your state. Put simply, our survey respondent's unobserved tastes are not random this month if they chose to purchase an iPhone 5 months ago. Properly drawing from the conditional distribution would be intractable, and imposing that the distribution of random coefficients is state-independent would be unrealistic.

Second, we do not directly observe switching in the dataset. If I observe a survey respondent who purchased an iPhone this month, I do not know what their state was when they arrived in this decision period: they may have been on contract or not, and they may have had a smartphone or not. One approach to measure the likelihood of this observation would be to look at the empirical distribution of states from the previous month for the given market and determine the likelihood of observing an individual purchase an iPhone this month, given the distribution of states in the previous month. This is feasible, although computationally costly, and relies heavily on the quality of the survey sample from that particular market.

Finally, direct estimation of each survey respondent would involve maximizing a likelihood over approximately 600,000 observations, a non-trivial task. Including

random coefficients would increase the computational burden linearly in the number of simulation draws per individual. Even if we were to ignore state-dependence and match aggregate market-level shares for each market and each month, the sample noise is problematic, particularly in smaller markets, and leads to cases of zero shares for some handset-network bundles, whose likelihood is undefined.

Taken together, this is evidence that this dataset does not lend itself to direct estimation and that serial correlation of tastes is an important aspect of this market to capture. For these reasons, I proceed with the model described in Section 4.

B.4 Bias-Corrected Objective Function and Inference

The bias-corrected objective function arises form the fact that, as has been noted before, the objective function

$$Q_{LNS}^{naive}(\theta) = \frac{1}{L} \sum_{l=1}^{L} \left\{ \left(\psi_l^0 - \psi_l^{NS}(\theta) \right)^2 \right\}$$

where moments are indexed by $l = 1..L$ results in a biased estimate when minimized. This is because minimizing the above has as its first order condition

$$H(\theta) \equiv \sum_{l=1}^{L} \left\{ \left(\psi_l^0 - \psi_l^{NS}(\theta) \right) \frac{\partial \psi_l^{NS}(\theta)}{\partial \theta} \right\} = 0$$

which, at the true value θ^0, has a non-zero expectation due to correlation between the simulated moment and its derivative; specifically,

$$H(\theta^0) = -E \left[Var \left(\psi^{NS}(\theta^0) \right) \right]$$

The bias-corrected objective function obtains a consistent estimate of this above

covariance and subtracts it from the naive objective function, resulting in a consistent estimator.

Confidence intervals are obtained using suggestions from Laffont, Ossard & Vuong (1995). Proposition 3 of the former paper establishes a method of estimating confidence intervals that correct for simulation bias (see pp. 964 for estimating equations). I use this suggestion in the construction of the confidence intervals for the point estimates of the parameters. For the confidence intervals of the counterfactuals, I bootstrap 200 draws from the estimated parameter distribution and report the 5th and 95th percentiles of the estimates.[1]

B.5 Robustness

One attractive feature of this setting is that carriers are not permitted to charge different prices in different markets. With 90 markets of data, I therefore have prices set at a national level but market-level variation in terms of the product quality (dropped calls). Since price is fixed across markets, I do not need to be concerned about price being correlated with market-level variation in products. However, since carriers are not able to vary prices across markets, it is likely that they may vary other factors in response to differences in their product quality in a given market. It is for this reason that I explicitly include a carrier's share of advertising spend in the demand for a "flagship" handset. Another concern may be a carrier's retail presence: I regressed the share of a carrier's customers in a market who reported that they purchased their device from one of the carrier's own retail stores (as opposed to a national chain or online) on the carrier's network quality and found no relationship in the data. This

---

[1] For counterfactuals that involve re-computing the price equilibrium, I cannot confirm that the bootstrap method is valid, as I cannot prove that iterating best responses leads to a unique price equilibrium in this model.

leads me to conclude that carriers are not significantly altering their retail presence in response to their network quality.

B.6 Exogeneity of Network Quality

In this section I will argue that network quality is exogenous and that any potential bias would work against by results under mild assumptions.

First, as shown in Figure 3.1, network quality does not vary much over time in the data. This is due to the fact that is it difficult for carriers to radically improve their network quality. Erecting new cell sites requires a long permitting process that varies by city and county, and even with sufficient spectrum holdings, it is a challenging engineering task to construct a high performance wireless network. For example, AT&T has the largest specturm holdings of any wireless carrier, but does not have the highest quality network.[2] The fact that network quality varies at all across markets is testament to the fact that, while every carrier would like to have high network quality in every market, there are exogenous factors that affect the quality of a carrier's network across markets.

Second, a possible source of unobserved demand shocks that could be correlated with a carrier's network quality in a market is the availability of "bundled services", where consumers purchase wireless service in conjunction with any of home television, internet, or landline services and a bundle discount. The survey data contains a question about bundled services, which I use to contruct an indicator variable for markets in which Verizon and AT&T offer such bundles. The concern would be that this may increase demand, and that carriers may invest differently in network qual-

---

[2] Sprint Nextel Corporation, "Petition to Deny", briefing filed in the application of AT&T Inc. and Deutsche Telekom AG.
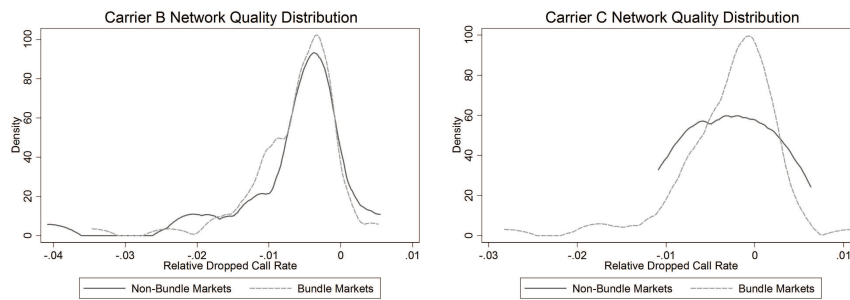
Figure B.3: Network Quality in "Bundle" and "Non-Bundle" Markets

ity in such markets. I perform a t-test for each of those carriers to see if the mean network quality in "bundle" and "non-bundle" markets differ, and fail to reject the null hypothesis that the means are identical (I get the same result using a single month's network quality and using the average network quality over all 26 months). Below are non-parametric density plots of each carrier's network quality (relative to market average) for "bundle" and "non-bundle" markets for Survey Month 40 in the data. The plot for Carrier B shows very similar distributions, and while the plot for Carrier C shows less similar distributions, there does not seem to be a systematic difference. I conclude from this that offering bundled services is uncorrelated with network quality.

Finally, I will argue that any possible bias is likely to work against my results. If carriers invest *less* in markets where they have positive demand shocks, then my estimate of the tastes for network quality would be biased towards zero, which would work against my findings in Counterfactual 1. It would in fact be optimal for a carrier to invest less in such markets if a positive demand shock reduces the marginal return on investment. This is likely to be the case whenever there are diminishing returns to network quality, a reasonable assumption. Even if a carrier perceived constant returns

in network quality, this finding would still hold as long as a carrier's cost function to achieve a given level of network quality were convex, also a reasonable assumption.

C. APPENDIX TO CHAPTER 4

C.1   *Robustness*

In appendix table C.1, we show how our key results vary with alternative specifications of the model. The columns of the table show (1) share of multi-paper markets that are diverse in our baseline model, (2) share of multi-paper markets that are diverse when firms ignore their competitors, (3) share of all markets that are diverse in our baseline model, and (4) share of all markets that are diverse when firms form joint operating agreements.

The first row of the table repeats the results from our main specifications for reference.

The second and third rows explore the sensitivity of our findings to the calibrated value of marginal costs we use, increasing and decreasing the marginal costs by 10 percent relative to the baseline value and re-estimating the model.

The fourth and fifth rows explore the sensitivity of our findings to the calibrated value of a_h we use, increasing and decreasing a_h by 10 percent relative to the baseline value and re-estimating the model.

The sixth row presents estimates from a specification in which we modify the demand model to treat the number of firms available in a town as endogenous. In particular, we model the number of firms J_t in a town t as a Poisson random variable whose log mean is a linear function of $log(S_t)$, ρ_t, ρ_t^2.

The seventh row adds flexibility to the fixed cost distribution in the supply model by allowing $\frac{\kappa_m}{S_m}$ to be distributed logistic with mean $\mu_\kappa^0 + \mu_\kappa^1 log\,(S_m) + \mu_\kappa^2 log(S_m)^2$.

The eighth row presents estimates from a specification in which we allow greater flexibility in the way in which consumer ideology affects the affiliations of newspapers that are available in a given town. In particular, we assume that for each newspaper j available in town t,

$$\Pr\left(\tau_j = R\right) = \text{logit}^{-1}\left(\mu_\rho^0 + \mu_\rho^1 \text{logit}\,(\rho_t) + \mu_\rho^2 \text{logit}\,(\rho_t)^2\right).$$

The ninth row tightens the population restrictions defining the universe of potential daily newspaper markets by 25%. This is done by dropping all market pairs containing a market with population smaller than 3,750 or larger than 75,000.

The tenth row presents estimates from a subsample of the data in which any market pair containing one or more independent newspapers as of 1924 is excluded.

The eleventh row presents estimates from a subsample of the data in which any market pair containing one or more unaffiliated newspapers as of 1924 is excluded.

The twelfth row presents estimates from a subsample of the data in which we exclude any market pair containing a market within 100km of any of the ten most populous cities as of the 1920 Census.

The thirteenth row presents estimates from a subsample of the data in which we drop town pairs for which our town-level circulation data omit a newspaper in at least one town's nearest news market.

The fourteenth row presents estimates from a subsample of the data in which any market pair containing a market in the South is excluded. Because of the dominance of the Democratic party in the South, excluding markets in the South increases estimated diversity at baseline and in all counterfactuals, but the differences between

counterfactuals remain similar to our preferred estimate from the full sample.

The fifteenth row presents estimates from a subsample of the data which removes any market pair containing a pair of papers in different markets that are owned by the same chain as of 1932. (Our ownership data are from the 1932 *Editor and Publisher Yearbook*. The earlier annual directories that we use to construct our main sample do not include lists of chain-owned newspapers.)

The sixteenth row presents estimates from an alternate sample in which we include any town that is itself the headquarters of a daily newspaper.

Table C.1: Robustness checks

	% of multi-paper markets w/ diverse papers		% of news markets w/ diverse papers	
	Baseline	Ignore competitors	Baseline	JOAs
(1) Preferred estimate	0.59	0.34	0.15	0.30
Changing Calibrated Values				
(2) Increase marginal cost by 10%	0.59	0.32	0.17	0.33
(3) Decrease marginal cost by 10%	0.59	0.32	0.15	0.29
(4) Increase a_h by 10%	0.60	0.36	0.18	0.32
(5) Decrease a_h by 10%	0.58	0.30	0.15	0.33
Modifying Model Specification				
(6) Endogenous J in demand model	0.58	0.34	0.15	0.30
(7) Add flexibility to fixed cost distribution	0.87	0.46	0.87	0.85
(8) Add flexibility to affiliation choice in demand	0.58	0.34	0.15	0.30
Modifying Estimation Sample				
(9) Tighten population cut-offs for markets	0.50	0.29	0.14	0.30
(10) Remove markets with independent papers	0.55	0.31	0.15	0.29
(11) Remove markets with unaffiliated papers	0.55	0.30	0.14	0.28
(12) Remove markets near major cities	0.53	0.28	0.13	0.27
(13) Remove towns if missing data for nearby papers	0.60	0.28	0.16	0.33
(14) Remove markets in the South	0.60	0.34	0.17	0.32
(15) Remove pairs with cross-market co-ownership	0.55	0.33	0.14	0.29
(16) Include daily paper headquarter towns	0.62	0.29	0.16	0.29

Notes: See appendix for details.

134

BIBLIOGRAPHY

ACKERBERG, D., AND M. RYSMAN (2005): "Unobservable Product Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects," *RAND Journal of Economics*, 36, 771–788.

AGHION, P., AND P. BOLTON (1987): "Contracts as a Barrier to Entry," *American Economic Review*, 77, 388–401.

ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of Political Economy*, 113(1), 151–184.

AMBRUS, A., AND M. REISINGER (2006): "Exclusive vs overlapping viewers in media markets," Harvard University Mimeograph.

ANDERSON, S. P., YSTEIN FOROS, AND H. J. KIND (2011): "Competition for advertisers in media markets," University of Virginia Mimeograph.

ARGENTESI, E., AND L. FILISTRUCCHI (2007): "Estimating market power in a two-sided market: The case of newspapers," *Journal of Applied Econometrics*, 22(7), 1247–1266.

ARMSTRONG, M. (2002): "Competition in two-sided markets," Nuffield College Mimeograph.

ASKER, J. (2005): "Diagnosing Foreclosure due to Exclusive Dealing," Leonard N. Stern School of Business, NYU mimeo.

BALDASTY, G. J. (1992): *The Commercialization of News in the Nineteenth Century*. University of Wisconsin Press, Madison, WI.

——— (1999): *E.W. Scripps and the Business of Newspapers*. University of Illinois Press, Urbana, IL.

BECKER, G. S. (1958): "Competition and democracy," *Journal of Law and Economics*, 1, 105–109.

BERNHEIM, B. D., AND M. D. WHINSTON (1998): "Exclusive Dealing," *The Journal of Political Economy*, 106(1), 64–103.

BERRY, S., AND A. PAKES (2007): "The Pure Characteristics Demand Model," *International Economics Review*, 48(4), 1193–1225, Special Issue Economics to Econometrics: Contributions in Honor of Daniel L. McFadden, Charles F. Manski and Whitney Newey, eds.

BERRY, S. T., AND J. WALDFOGEL (2010): "Quality and market size," *Journal of Industrial Economics*, 58(1), 1–31.

BESANKO, D., AND M. K. PERRY (1994): "Exclusive dealing in a spatial model of retail competition," *International Journal of Industrial Organization*, 12(3), 297 – 329.

BORK, R. H. (1978): *The Antitrust Paradox: A Policy at War with Itself*. Basic Books, New York.

BRESNAHAN, T. F., AND P. C. REISS (1991): "Entry and competition in concentrated markets," *Journal of Political Economy*, 99(5), 977–1009.

BULOW, J., J. GEANAKOPLOS, AND P. KLEMPERER (1985): "Multimarket oligopoly: strategic substitutes and strategic complements," *Journal of Political Economy*, 93, 488–511.

BUSTERNA, J. C., AND R. G. PICARD (1993): *Joint Operating Agreements: The Newspaper Preservation Act and its Application*. Ablex Publishing Corporation, Norwood, NJ.

BYRD, R. H., J. NOCEDAL, AND R. A. WALTZ (2006): *Large-Scale Nonlinear Optimization* chap. KNITRO: An integrated package for nonlinear optimization. Springer-Verlag, New York.

CHANDRA, A., AND A. COLLARD-WEXLER (2009): "Mergers in two-sided markets: An application to the Canadian newspaper industry," *Journal of Economics & Management Strategy*, 18(4), 1045–1070.

CHERNOZHUKOV, V., AND H. HONG (2003): "An MCMC approach to classical estimation," *Journal of Econometrics*, 115, 293–346.

CHIANG, C.-F. (2010): "Political differentiation in newspaper markets," National Taiwan University Mimeograph.

DIXIT, A. K., AND J. E. STIGLITZ (1977): "Monopolistic competition and optimum product diversity," *American Economic Review*, 67(3), 297–308.

DOWNS, A. (1957): "An economic theory of political action in a democracy," *Journal of Political Economy*, 65(2), 135–150.

DRAGANSKA, M., M. MAZZEO, AND K. SEIM (2009): "Beyond plain vanilla: Modeling joint product assortment and pricing decisions," *Quantitative Marketing and Economics*, 7(2), 105–146.

DURANTE, R., AND B. KNIGHT (Forthcoming): "Partisan control, media bias, and viewer responses: Evidence from Berlusconi's Italy," *Journal of the European Economic Association*.

EINAV, L. (2007): "Seasonality in the US motion picture industry," *The RAND Journal of Economics*, 38(1), 127–145.

———— (2010): "Not all rivals look alike: Estimating an equilibrium model of the release date timing game," *Economic Inquiry*, 48(2), 369–390.

FAN, Y. (2010): "Ownership consolidation and product quality: A study of the us daily newspaper market," Working Paper, University of Michigan.

FCC (2010): "Notice of inquiry. FCC 10-92.," .

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

GENTZKOW, M., E. L. GLAESER, AND C. GOLDIN (2006): *Corruption and Reform: Lessons from America's Economic History*chap. The rise of the fourth estate: How newspapers became informative and why it mattered, pp. 187–230. University of Chicago Press, Chicago, IL.

GENTZKOW, M., N. PETEK, J. M. SHAPIRO, AND M. SINKINSON (2011): "Political influence on US newspapers," Poster Presented at the Midwest Political Science Association Meeting March 31, 2011, Chicago IL.

GENTZKOW, M., AND J. M. SHAPIRO (2010): "What drives media slant? Evidence from US daily newspapers," *Econometrica*, 78(1), 35–71.

GENTZKOW, M., J. M. SHAPIRO, AND M. SINKINSON (2011): "The effect of newspaper entry and exit on electoral politics," *American Economic Review*, 101(7), 2980–3018.

GEWEKE, J. F., AND M. P. KEANE (1996): "Bayesian Inference for Dynamic Discrete Choice Models Without the Need for Dynamic Programming," Federal Reserve Bank of Minneapolis Working Paper 564.

GOURIEROUX, C., AND A. MONFORT (1993): "Simulation-based inference," *Journal of Econometrics*, 59, 5–33.

GOWRISANKARAN, G., AND M. RYSMAN (2011): "Dynamics of Consumer Demand for New Durable Goods," mimeo.

GROSECLOSE, T., AND J. MILYO (2005): "A measure of media bias," *Quarterly Journal of Economics*, 120(4), 1191–1237.

GUSTAFFSON, K. E., H. RNEBRING, AND D. A. L. LEVY (2009): "Press subsidies and local news: The Swedish case," Reuters Institute for the Study of Journalism Working Paper.

HAMILTON, J. T. (2006): *All the News That's Fit to Sell: How the Market Transforms Information into News*. Princeton University Press, Princeton, NJ.

HECKMAN, J. J. (1981): *Structural Analysis of Discrete Data with Econometric Applications*chap. The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, pp. 179–195. MIT Press, Cambridge, MA.

HEISS, F., AND V. WINSCHEL (2008): "Likelihood approximation by numerical integration on sparse grids," *Journal of Econometrics*, 144(1), 62–80.

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(2), 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), 1–28.

KAPLAN, R. L. (2002): *Politics and the American Press: The Rise of Objectivity, 1865-1920*. Cambridge University Press, Cambridge.

KATZ, M. L. (1989): "Chapter 11 Vertical contractual relations," in *Handbook of Industrial Organization*, ed. by R. Schmalensee, and R. Willig, vol. 1, pp. 655 – 721. Elsevier.

KIELBOWICZ, R. B. (1994): "Cost accounting in the service of policy reform: Postal rate making," *Social Science Quarterly*, 75(2), 284–299.

LAFFONT, J.-J., H. OSSARD, AND Q. VUONG (1995): "Econometrics of First-Price Auctions," *Econometrica*, 63(4), 953–980.

LARCINESE, V., R. PUGLISI, AND J. M. SNYDER (2007): "Partisan bias in economic news: Evidence on the agenda-setting behavior of US newspapers," *Journal of Public Economics*, 95, 1178–1189.

LEE, R. (2010): "Vertical Integration and Exclusivity in Platform and Two-Sided Markets," New York University mimeo.

MANKIW, N. G., AND M. D. WHINSTON (1986): "Free entry and social inefficiency," *RAND Journal of Economics*, 17(1), 48–58.

Mazzeo, M. J. (2002): "Product choice and oligopoly market structure," *RAND Journal of Economics*, 33(2), 221–242.

McFadden, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models," *Econometrica*, 57, 995–1026.

Mor, J. J., B. S. Garbow, and K. E. Hillstrom. (1980): *User Guide for MINPACK-1. Argonne National Laboratory Report ANL-80-74.* Argonne National Laboratory, Argonne, IL.

Mullainathan, S., and A. Shleifer (2005): "The market for new," *American Economic Review*, 95(4), 1031–1053.

Murphy, K. M., and R. H. Topel (1985): "Estimation and inference in two-step econometric models," *Journal of Business and Economic Statistics*, 3(4), 370–379.

——— (1990): *Advances in the Theory and Measurement of Unemployment*chap. Efficiency wages reconsidered: Theory and evidence. St. Martin's Press, New York.

Murschetz, P. (1998): "State support for the daily press in Europe: A critical appraisal," *European Journal of Communication*, 13(3), 291–313.

of US, C. U. (2009): "CELL PHONES: Our tests of 70 standard and smart models show they're sharing many more features," *Consumer Reports*, 74(1), Albany, NY.

Pakes, A. (1986): "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54(4), 755–784.

Pakes, A., and D. Pollard (1989): "Simulation and the asymptotics of optimization estimators," *Econometrica*, 57(5), 1027–1057.

Petrova, M. (2009): "Newspapers and parties: How advertising revenues created an independent press," CEFIR/NES Working Paper No. 131.

Posner, R. (1976): *Antitrust law: An economic perspective*. University of Chicago Press, Chicago.

Powell, M. J. D. (1970): *A hybrid method for nonlinear equations*. Gordon and Breach, London.

Puglisi, R. (2011): "Being the New York Times: The political behaviour of a newspaper," *The B.E. Journal of Economic Analysis & Policy*, 11(1).

Puglisi, R., and J. M. Snyder (2008): "Media coverage of political scandals," NBER Working Papers 14598.

Qian, N., and D. Yanagizawa (2010): "Watchdog or lapdog? Media and the US government," NBER Working Paper w15738.

Reiss, P. C., and P. T. Spiller (1989): "Competition and entry in small airline markets," *Journal of Law & Economics*, 32, S179.

Rey, P., and J. Stiglitz (1995): "The role of exclusive territories in producers' competition," *RAND Journal of Economics*, 26(3), 431–451.

Rysman, M. (2004): "Competition between networks: A study of the market for yellow pages," *The Review of Economic Studies*, 71(2), 483–512.

Segal, I., and M. D. Whinston (2000): "Exclusive contracts and protection of investments," *RAND Journal of Economics*, 31, 603–633.

Seim, K., and J. Waldfogel (2010): "Public monopoly and economic efficiency: Evidence from the Pennsylvania Liquor Control Boards entry decisions," NBER Working Paper Series No. 16258.

Skrainka, B. S., and K. L. Judd (2011): "High performance quadrature rules: How numerical integration affects a popular model of product differentiation," CeMMAP Working Paper CWP03/11.

Sweeting, A. (2010): "The effects of horizontal mergers on product positioning: Evidence from the music radio industry," *RAND Journal of Economics*, 41(2), 372–397.

Whinston, M. D. (2006): *Lectures on Antitrust Economics*. MIT Press, Cambridge, MA.

Wilbur, K. C. (2008): "A two-sided, empirical model of television advertising and viewing markets," *Marketing Science*, 27(3), 356–378.

Yewdall, C. (1928): *Report of the Cost Finding Committee*. Inland Daily Press Association, Des Plaines, IN.

Zhu, T., H. Liu, and P. K. Chintagunta (2011): "Wireless Carriers Exclusive Handset Arrangements: An Empirical Look at the iPhone," NET Institute working paper.