

NBER WORKING PAPER SERIES

EVALUATING WIRELESS CARRIER CONSOLIDATION
USING SEMIPARAMETRIC DEMAND ESTIMATION

Patrick Bajari
Jeremy T. Fox
Stephen Ryan

Working Paper 12425
<http://www.nber.org/papers/w12425>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2006

Bajari thanks the National Science Foundation for generous research support. Fox thanks the NET Institute and the Olin Foundation for generous funding. Thanks to helpful comments from seminar participants at the International Industrial Organization Conference (IIOC), the Northwestern and University of Chicago IO-Marketing conference, and the STIET seminar at the University of Michigan. We thank Austan Goolsbee, Ali Hortacsu and Chad Syverson for providing access to the Forrester Data, and for helpful comments. We thank Dennis Carlton, Anindya Ghose and Bill Rogerson for detailed conversations. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2006 by Patrick Bajari, Jeremy T. Fox and Stephen Ryan. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Evaluating Wireless Carrier Consolidation Using Semiparametric Demand Estimation
Patrick Bajari, Jeremy T. Fox and Stephen Ryan
NBER Working Paper No. 12425
August 2006
JEL No. L1, C1

ABSTRACT

The US mobile phone service industry has dramatically consolidated over the last two decades. One justification for consolidation is that merged firms can provide consumers with larger coverage areas at lower costs. We estimate the willingness to pay for national coverage to evaluate this motivation for past consolidation. As market level quantity data is not publicly available, we devise an econometric procedure that allows us to estimate the willingness to pay using market share ranks collected from a popular online retailer, Amazon. Our semiparametric maximum score estimator controls for consumers' heterogeneous preferences for carriers, handsets and minutes of calling time. We find that national coverage is strongly valued by consumers, providing an efficiency justification for across-market mergers. The methods we propose can estimate demand for other products using data from Amazon or other online retailers where quantities are not observed, but product ranks are observed. Since Amazon data can easily be gathered by researchers, these methods may be useful for the analysis of other product markets where high quality data are not publicly available.

Patrick Bajari
Professor of Economics
University of Minnesota
1035 Heller Hall
271 19th Avenue South
Minneapolis, MN 55455
and NBER
bajari@econ.umn.edu

Jeremy T. Fox
University of Chicago
Department of Economics
1126 E. 59th Street
Chicago, IL 60637
fox@uchicago.edu

1 Introduction

Currently the US mobile phone service industry is dominated by four large, national carriers: Cingular, Sprint Nextel, T-Mobile and Verizon. The emergence of national carriers is a recent phenomenon. In 1988, the company serving the largest number of the top 20 markets was US West, which served only four markets.

Major mergers in the United States must be approved by the federal government. Companies proposing mergers typically argue that the merger will enhance consumer welfare through better products. Demand estimation can cast light on whether improved features provide a quantitatively grounded motivation for mergers to be in the public interest. A major measure of product quality for a wireless carrier is the extent of its geographic coverage. This paper examines an industry with \$113 billion in annual revenue after a series of very large mergers to estimate consumers' value for national coverage.

Academic work on wireless mergers has been hampered by a lack of quantity data. In the mobile phone industry, the relevant unit of analysis for demand estimation is a geographic market, such as a metropolitan area. The competition for subscribers is market specific, as federal regulations prohibit a carrier that does not have one of a scarce number of geographic market-specific licenses from signing up new customers in that market. Mobile phone carriers release data only on national customer counts, not the numbers of customers per geographic market.

One source that does distribute publicly available share data is the online retailer Amazon.com. Amazon sells more than a billion dollars worth of products a year, including mobile phones and subscription plans. Along with a comprehensive enumeration of all characteristics and features for each plan, Amazon lists each plan's sales rank over a recent time interval. We use these market share ranks as our dependent variables in estimating the willingness to pay for various plan characteristics. As Amazon is a retailer helping customers choose between plans, we can see all the subscription plan characteristics that customers observe on the site. Therefore, we use Amazon to collect rich data on both the dependent variables, market-specific share ranks, and the independent variables, plan characteristics, including prices.

Using data on market share ranks introduces an estimation challenge. Most previous demand estimation methods for market-level data require continuous measures of shares, even if the choice of a single consumer is a discrete choice between a set of competing products. Exceptions include Brynjolfsson, Hu and Smith (2003), Chevalier and Goolsbee (2003), and Ghose and Sundararajan (2006), who use insider knowledge and online experiments to empirically verify that book sales on Amazon follow a power law. These studies then use the power law distribution to interpolate continuous measures of market shares. However, interpolation only works if one can use other sources to estimate the distribution of continuous market shares from market share ranks. And since a consumer can only choose between 70–80 mobile phone subscription plans, many fewer than the millions of books on Amazon, the same power law that holds for books is unlikely to hold for calling plans.

Existing methods for estimating aggregate data discrete choice models impose a parametric structure on demand that makes precise predictions about market shares, given a trial vector of parameters. For example, least squares can be used to estimate demand parameters when each consumer maximizes his utility according to the standard logit random utility model. The known distribution of the error terms in the logit model generates precise predictions about market shares. Many recent papers in industrial organization and related fields have expanded on these methods, including Chintagunta, Jain and Vilcassim (1991), Berry (1994), Berry, Levinsohn and Pakes (1995), Allenby and Rossi (1999), Nevo (2001), Petrin (2002) and Bajari and Benkard (2005).

Data on market share ranks are less precise than market share levels, and cannot make use of the precise

predictions arising from assuming a distribution for the errors. Our approach therefore belongs to the class of semiparametric discrete choice estimators, which do not rely on assuming the error term distribution is known. Assuming certain broad properties about the error terms' distribution, Manski (1975) shows that a consumer's probabilities of making discrete choices are rank ordered by the mean payoffs of the choices. Our insight is that by aggregating the independent decisions of many such consumers in a market, the resulting market shares are rank ordered by the mean payoffs of each product. We adapt Manski's identification strategy and the resulting maximum score estimator to the case of market-level data, which obviates the need to interpolate market shares from ranks.

A serious concern with recovering the willingness to pay for plan characteristics from aggregate data is that the aggregate ranks may reflect several sources of unobserved heterogeneity. Our approach to dealing with this problem is to compare the ranks of plans that are as close to each other as possible with the exception of the characteristic of interest. In the language of the nested logit, we make pairwise comparisons across products in the same product nest. This strategy is nonparametric: we difference out unobserved heterogeneity without imposing assumptions about the form of the unobserved heterogeneity across nests. In the absence of sampling error, mapping from the variation in the data to the observed estimates is transparent: between two comparable plans, the one with more market share has a higher mean payoff.

In our application, the parameter with the most interest from an antitrust perspective is the willingness to pay for national coverage. To estimate the willingness to pay for this characteristic, we compare plans offering regional coverage to those offering national coverage with similar numbers of monthly calling minutes from the same carrier. Our approach controls for a carrier's quality, a customer's idiosyncratic preferences for each carrier's phones, a customer's view of each carrier's coverage around his or her home and workplace, as well as a customer's willingness to pay for minutes of calling time. Our estimator is consistent under these consumer and carrier specific fixed effects, which may be correlated with price and included plan characteristics.

The use of Amazon data introduces sample selection issues because Amazon does not offer plans from all carriers. Typically, a discrete choice estimator such as the random coefficients logit is inconsistent if data on some choices are missing. However, Fox (2006) has shown the general class of estimators that our approach belongs to are consistent when the econometrician observes a subset of the choices available both on and offline, which is the case with Amazon data.

Another possibility is that the willingness to pay for consumers who buy phones on Amazon differs from the willingness to pay for consumers who shop in brick and mortar stores. Our results are less attractive for antitrust analysis if our data do not cover the customer population that merger activity targets. We examine the representativeness of Amazon customers by referring to the Forrester survey of American households. We find that mobile phone users that shop on Amazon tend to be wealthier and younger than customers that do not shop on Amazon. They also travel more, which suggests Amazon users value national coverage more than non-users. Therefore, our estimate of the value of national coverage may be an upper bound.

We evaluate a motivation for mobile phone carrier mergers. We do not model a carrier's choice of its menu of plans. Recent work has considered the menu of products offered by a monopolist, including the three-part tariff structure (monthly fee, bucket of included minutes, overage charge) of subscription plans (Miravete and Röller, 2004; Seim and Viard, 2004; Crawford and Shum, 2006). Competing by offering a menu of plans may alter the calculations behind mergers, but most press coverage focuses on broad efficiency and market power explanations for mergers.

Our data allow us to bound the willingness to pay for national coverage. Our main result is that the willingness to pay for national coverage is between 0.65 and 0.77 cents a minute, corresponding to between 7.3% and 8.6% of the monthly cost of the most popular calling plans. Extrapolating, the 208 million cellular phone users in the United States in 2005 have a total annual willingness to pay of \$9.1 to \$10.7 billion. The large willingness to pay shows that consolidation that reduces the interfirm contractual costs of providing national coverage is a quantitatively grounded motivation for wireless consolidation.

The methods that we propose are useful for analyzing demand in other markets where product ranks are observable and quantities are unobservable. Many online retailers, including the giants Amazon and Wal-Mart, allow the user to sort alternative products from the most to the least popular. All else held constant, it is of course preferable to use data with actual quantities since weaker econometric assumptions can be used. Despite this disadvantage, we believe in many cases data from online retailers have some strengths compared to alternative data sources. First, in an online market, the economist is able to observe the exact information about a product presented by the retailer to the consumer. In many empirical studies of differentiated product markets, by comparison, there is a large gap between the economist's and the consumer's information about a product. Commonly, the economist only observes an average (or quantity weighted) price and a fairly incomplete list of product attributes. In online markets, the exact price is observed and online retailers often generate web pages in a manner that allows the consumers to compare a large number of characteristics across products. In our Amazon data, for example, it is possible to construct a matrix of 13 product characteristics for 70-80 plans across 22 markets. Second, measurement error may be less problematic in data collected from online retailers. The web pages we downloaded from Amazon are contracts that describe the product and terms offered by the retailer. Online retailers have strong incentives to make sure that such information is reported accurately. By comparison, measurement error is a common (and ignored) problem in many studies of differentiated product markets. Third, online data can be collected freely for a very diverse set of product offerings. Leading online retailers such as Amazon and Wal-Mart have product offerings in a mind boggling array of categories. In many of these categories, high quality data are not publicly available to researchers from other sources.

2 Wireless Carrier Consolidation

Fox (2005) presents an overview of the history of mobile phone consolidation. The early history of consolidation featured only across-market mergers, as within market mergers were not allowed. The Federal Communications Commission (FCC) is the primary regulator of the mobile phone service industry. Starting in 1996, the FCC began a gradual process of loosening and finally eliminating the "spectrum cap", the fraction of the public radio waves allocated for mobile phone use that an individual carrier can control in a given geographic market. As governmental objections to mergers have declined, more across and within-market mergers have occurred. As we address in this paper, market-specific subscriber numbers are not available publicly. Therefore, it is difficult to assess carrier market power at the level of a particular city using standard data sources.¹

Nevertheless, the high, and relatively recent, degree of concentration in the mobile phone service industry raises the warning flag that carriers may be exploiting retail market power to keep prices high. FCC Commis-

¹The FCC uses data on the number of telephone numbers assigned to carriers to approximate market shares, rather than using data on actual customers. The FCC writes in response to a Department of Justice request to access its data on market shares, "The Commission has recognized that disaggregated, carrier-specific forecast and utilization data should be treated as confidential and should be exempt from public disclosure under 5 U.S.C. Section 552(b)(4)."

sioner Michael J. Copps, who has access to restricted-access market share estimates, writes in his statement of approval for the Sprint / Nextel merger the warning, “The average US market’s HHI score has grown from 2,900 (before the Cingular/AT&T merger) to 3,100 (after the Cingular/AT&T merger) to 3,300 (after the Sprint/Nextel merger).” Copps correctly writes that an HHI (Herfindahl-Hirschman Index) of 3,300 implies that there are now an equivalent of three equal sized competitors in most markets. In other industries, the Department of Justice’s horizontal merger guidelines suggest that any industry with an HHI about 1,800 is “highly concentrated.”

To merge, two carriers must win approval from the FCC and the Department of Justice. The FCC writes in its approval of the Sprint/Nextel merger that a merger must “serve the public interest, convenience, and necessity.” Carriers promoting the regulatory approval of their mergers often write of the welfare-enhancing benefits of such mergers. Some of these benefits accrue on the cost side, and are invisible to consumers. Other benefits appear on both the cost and demand sides, and some mainly help consumers. Our goal in this paper is to use the revealed preferences of consumers to evaluate whether one of the claims made by wireless carriers is quantitatively important in terms of customers’ willingnesses to pay for new and improved features.

Most carriers in the United States offer their customers national calling areas through a combination of each carrier’s native coverage and roaming agreements signed with other carriers. Data on the early cellular industry show that the per-minute customer charge for roaming was around 55% higher than the charge for placing calls in a customer’s native coverage area (Fox, 2005). Today, many subscription plans make roaming charges invisible to customers. However, published industry comments suggest that carriers transfer high per-minute fees between each other when one carrier’s customer travels to a roaming market and places a call.²

There are both technological and incentive reasons for why a merged carrier might be able to achieve lower costs and greater benefits in providing national coverage. On the technology side, the integration of features such as voice mail notifications, internet access and push-to-talk is complex. Even if voice calls work while roaming, more advanced features may not, or the more advanced features may be costly to provide. A merger can standardize the implementation of advanced features across the entire carrier’s coverage area.

On the incentive side, carriers seem reluctant to provide roaming customers from other carriers access to some of these advanced features. For example, SouthernLinc and Nextel both use the same technology for standard voice as well as push-to-talk calls. Nextel has a much larger coverage area, and allows SouthernLinc roamers the right to place standard voice but not push-to-talk calls. A related issue is pricing: the carrier charging roaming fees can only indirectly (through the roaming fee) affect how the home carrier of a traveler passes along the charge. As a consequence, standard models of double marginalization predict that the sum of the profits of the roaming and home carriers will be less than the profits of a merged firm.

It is a theory of the firm issue why non-merged carriers cannot agree to charge roamers low per-minute rates. Whatever the reason, it appears correct that across-market mergers expand native calling areas and thus reduce roaming charges. If cost savings are passed on to consumers, lower roaming charges reduce the price premium national calling plans charge over regional plans where customers are directly billed for roaming. In the next section, we present an econometric methodology that can be used to measure consumers’ valuation for national calling.

²Some carriers have decided to take advantage of the network aspect of their products by offering free in-network calling. If two Verizon or two Cingular customers talk, the length of the call is not deducted from either customer’s bucket of included minutes from their subscription plans. Carriers merging create larger networks so that customers can better exploit free in-network calling. Unfortunately, major carriers either include unlimited in-network calling as part of all plans, or offer it as an add-on option. There is no variation within a carrier in whether in-network calling is included in a subscription plan, so we cannot estimate its value without using across-carrier variation in market share ranks.

3 Model and Estimator

3.1 Willingness to Pay

To fix ideas, suppose that the only observable characteristics of a plan are national coverage and price. We specify agent i 's utility for plan j , belonging to nest h in market m , as:

$$u_{ijhm} = x_{jm}^{\text{Nat}} \beta^{\text{Nat}} - p_{jm} + v_{ihm} + \xi_{jm} + \varepsilon_{ijm},$$

where x_{jm}^{Nat} is equal to 1 when a plan has national coverage. We discuss the error terms in detail below. We measure price in terms of cost per anytime minute, or

$$p_{jm} = \frac{\text{monthly price of plan } j \text{ in market } m}{\# \text{ of anytime calling minutes of plan } j \text{ in market } m}.$$

In this specification, national coverage shifts around the value of an anytime minute. With only regional coverage, an anytime minute has a normalized value of $-p_{jm}$. With national coverage, an anytime minute has a mean value of $x_{jm}^{\text{Nat}} \beta^{\text{Nat}} - p_{jm}$. The estimable parameter β^{Nat} is then a typical customer's per-minute willingness to pay for having no-surcharge national coverage instead of having to pay a surcharge when traveling. A motivation for approving mergers is that many customers have high valuations for national service. If so, β^{Nat} should be high.

We model national coverage as shifting around the value of a minute of talk time because eliminating surcharges for making calls is useful for every call. People who make more calls will have a higher total value for national coverage.

A consumer's utility for a subscription plan has the deterministic component $x\beta - p$. In consumer theory, any positive monotonic transformation of a utility function yields the same ordinal ranking of alternatives. Consumer choice data can be used to identify the tradeoffs from varying observable plan characteristics in the ordinal ranking of plans, not the underlying utility units each plan gives. We normalize the scale of utility to be in cents per minute of calling time. The coefficient on the cost per minute is -1 .

We will account for other plan characteristics. We generalize our specification to allow for a vector x_{jm} of d observable, non-price characteristics. Consequently, β becomes a vector of length d of willingness to pay parameters.

3.2 Customer Heterogeneity

Two observationally identical consumers may choose different plans. The failure of plan characteristics to explain all plan choices is captured by adding error terms to the utilities of choices. The error terms represent unobserved customer heterogeneity: payoffs about individual choices that are not observed by the econometrician. While agents have perfect information about their choices, to an econometrician the agents act probabilistically as the data do not contain all factors relevant to decision making.

We observe J_m plans in market m . We divide the J_m plans into H_m nests, which are groups of subscription plans. The utility u_{ijhm} of customer i for subscription plan j from nest h in geographic market m is

$$u_{ijhm} = x'_{jm} \beta - p_{jm} + v_{ihm} + \xi_{jm} + \varepsilon_{ijm},$$

where v_{ihm} is a customer and nest specific unobserved preference, ξ_{jm} is a market and subscription plan specific error, and ϵ_{ijm} is a customer and subscription plan specific error. Customer i chooses plan j when j maximizes his or her utility, meaning

$$u_{ijhm} > u_{iklm} \forall k \in J_M, k \neq j, \quad (1)$$

where l is the nest of product k , and ties occur with probability zero.

We use data on aggregate market share ranks, not on individual consumers. Therefore, we use this utility specification to form predictions about market share ranks for subscription plans. Markets share ranks are formed by aggregating the individual purchase decisions of the customers in the market to create market share levels, and then ordering the levels to produce ranks.

We use arrows to represent collections of items. Let \vec{x}_m be the $J_m \times (d+1)$ matrix of the observable characteristics of all plans in market m . Price is included in \vec{x}_m for notational convenience. For a consumer, \vec{v}_{im} is the vector of the H_m nest fixed effects.

We summarize our statistical assumptions about unknown error terms in the following assumption.

Assumption 1. *1. The consumer and nest errors v_{ihm} are fixed effects that can be correlated with each other and with any observable product characteristics, including price.*

2. For a given consumer i , each consumer and plan error term ϵ_{ijm} for a plan in the nest h is independent and identically distributed with the absolutely continuous density $f_h(\epsilon_{ijm} | \vec{x}_m, \vec{v}_{im})$ and has full support on the real line.

3. For a given market m , each plan and market error term ξ_{jm} for a plan in nest h is independent and identically distributed with the absolutely continuous density $g_h(\xi_{jm} | \vec{x}_m)$ and has full support on the real line.

We need to use Assumption 1 to prove that plans with higher mean payoffs are chosen more often, as our estimator maps market share ranks into mean payoffs. The following lemma is a slight generalization of a condition introduced for individual data in pioneering work by Manski (1975). Among two products in a nest and for a single consumer, the plan with the higher overall mean utility must have a higher probability of being chosen.

Lemma 1. *Let plans j and k be in the same nest h of plans in market m , and let consumer i be given. Under Assumption 1,*

$$x'_{jm}\beta - p_{jm} > x'_{km}\beta - p_{km}$$

if and only if

$$P(j | \beta, \vec{x}_m, \vec{v}_{im}) > P(k | \beta, \vec{x}_m, \vec{v}_{im}).$$

The proof is in an appendix. A choice probability $P(j | \beta, \vec{x}_m, \vec{v}_{im})$ reflects the uncertainty of an econometrician who does not observe the consumer and product specific errors ϵ_{ijm} and the product and market specific errors ξ_{jm} , but does know the structure of the decision model, the parameters β for the willingness to pay, the observed product characteristics including price \vec{x}_m , and the realizations of the consumer and carrier nest fixed effects \vec{v}_{im} .

Lemma 1 compares only two plans in the same nest h of plans. For simplicity, let h be the nest of all individual (non-family) plans offered by the carrier Verizon. Let customer i have an unobserved willingness to

pay of v_{im}^{Verizon} for Verizon, and likewise a willingness to pay of $v_{im}^{\text{T-Mobile}}$ for T-Mobile. We would require a functional form for the distribution of v_{im}^{Verizon} and $v_{im}^{\text{T-Mobile}}$ to predict the relative probability of picking a plan from Verizon and a plan from T-Mobile. Consider instead the decision of customer i to pick plan j over k , with both plans from Verizon. Customer i prefers j to k if

$$x'_{jm}\beta - p_{jm} + v_{im}^{\text{Verizon}} + \xi_{jm} + \varepsilon_{ijm} > x'_{km}\beta - p_{km} + v_{im}^{\text{Verizon}} + \xi_{km} + \varepsilon_{ikm}.$$

The common preference v_{im}^{Verizon} differences out, and the choice inequality reduces to the condition

$$x'_{jm}\beta - p_{jm} + \xi_{jm} + \varepsilon_{ijm} > x'_{km}\beta - p_{km} + \xi_{km} + \varepsilon_{ikm}.$$

By comparing two plans from Verizon, we do not need to make a functional form assumption about how preferences for Verizon differ across the population, and about whether the preferences for Verizon are correlated with preferences for plans from other carriers or with the observed characteristics of the plans (x_{jm}) from different carriers.

Nest and agent fixed effects v_{im} satisfy four roles in our application:

1. Each consumer may have a certain need to talk on the phone. If a nest is restricted to plans with similar numbers of minutes of airtime, then we control for a consumer's heterogeneous tastes for mobile communication.
2. Each carrier has its own network of cellular base stations, so one carrier may have better coverage near a consumer's home or workplace than another carrier. The fixed effects capture a consumer's views on the coverage of each carrier.
3. Consumers simultaneously buy calling plans and phones. Our data cannot pair individual phone and plan purchases. The fixed effects capture a consumer's views on each carrier's phones, as we define a nest narrowly enough so that all plans can be paired with the same set of phones.
4. Maybe Verizon optimally chose its phone lineup in conjunction with its menu of plans. Fixed effects can be correlated with observable plan characteristics, including price.

The additive separability of $x'_{jm}\beta - p_{jm}$ and v_{im}^{Verizon} is not mathematically necessary for our fixed effects strategy. In an ordered choice context, Abrevaya (2000) introduces a more general specification where the contributions of deterministic payoffs and fixed effects are weakly separable.

3.3 Market Share Ranks

We want to show that plans with greater payoffs will have higher market share ranks more often than not. Define the market share of product j in a market m with I_m consumers to be

$$s_{jm} = \frac{1}{I_m} \sum_{i=1}^{I_m} 1[i \text{ buys } j],$$

where $1 [i \text{ buys } j]$ is an indicator equal to 1 when consumer i buys subscription plan j . Because expectation and integration are linear operators, conditioning on the number of consumers I_m results in:

$$\begin{aligned} E [s_{jm} | I_m] &= E \left[\frac{1}{I_m} \sum_{i=1}^{I_m} 1 [i \text{ buys } j] | I_m \right] = \frac{1}{I_m} \sum_{i=1}^{I_m} E [i \text{ buys } j | I_m] = \\ &= \frac{1}{I_m} \sum_{i=1}^{I_m} E [E_{\epsilon, \xi} [1 [i \text{ buys } j] | \beta, \vec{x}_m, \vec{v}_{im}] | I_m] = \frac{1}{I_m} \sum_{i=1}^{I_m} E [P(j | \beta, \vec{x}_m, \vec{v}_{im}) | I_m], \end{aligned}$$

where the second-to-last equality is from the law of iterated expectations and the last equality uses the definition of a choice probability that integrates out consumer product specific error terms of the form ϵ_{ijm} and product specific error terms of the form ξ_{jm} . The additive separability of a market share allows us to sum the choice probabilities of consumers despite the presence of the same plan and market error ξ_{jm} in the decision of all consumers.

Consider two products j and k in the same nest h , and let $x'_{jm}\beta - p_{jm} > x'_{km}\beta - p_{km}$, where β is the true willingness to pay for plan characteristics in the data generating process. Under Assumption 1, Lemma 1 states that $P(j | \beta, \vec{x}_m, \vec{v}_{im}) > P(k | \beta, \vec{x}_m, \vec{v}_{im})$ for all consumers. By the above market share algebra,

$$E [s_{jm} | I_m] = \frac{1}{I_m} \sum_{i=1}^{I_m} E [P(j | \beta, \vec{x}_m, \vec{v}_{im}) | I_m] > \frac{1}{I_m} \sum_{i=1}^{I_m} E [P(k | \beta, \vec{x}_m, \vec{v}_{im}) | I_m] = E [s_{km} | I_m],$$

as each consumer chooses j more often than k . As $E [s_{jm} | I_m] > E [s_{km} | I_m]$ for any number of consumers I_m , $E [s_{jm}] > E [s_{km}]$ unconditionally. Even if an econometrician does not know the number of consumers in a market, a plan with a higher mean payoff than another plan in the same nest has a higher expected market share.

We work with market share rank data. For a given market, we see whether one plan has a higher market share than another plan. Let r_{jm} be the rank of plan j in market m , with higher ranks corresponding to plans with higher shares. If there are 70 plans, the plan with the highest market share has a market share rank of 70, not 1. We want to prove the property that a plan with a higher mean payoff will have higher market share rank more often than not. Given two random variables a and b , it is possible that a has a higher mean than b even though the random variable $1 [a > b]$ has a mean less than $1/2$. Fortunately, our decision model is well behaved and the expected market share ranks of products in the same nest are themselves rank ordered by their mean payoffs.

Lemma 2. *Let plans j and k be in the same nest h of plans in market m . Under Assumption 1,*

$$E [1 [r_{jm} > r_{km}] | \beta, \vec{x}_m] > E [1 [r_{jm} < r_{km}] | \beta, \vec{x}_m]$$

if and only if

$$x'_{jm}\beta - p_{jm} > x'_{km}\beta - p_{km}.$$

The proof of the lemma is cumbersome because the proof deals with both the integer nature of a market share rank and the different choice probabilities of consumers. See the appendix for details.

3.4 Set Identification Without Sampling Error

For this section only, we consider the special case of our model where product and market specific errors are not present, so that $\xi_{jm} = 0$ for all $j \in J_m$ and markets $m \in M$, where M is the set and number of all markets. We also assume that the number of consumers in each market m is large, so that $I_m \rightarrow \infty$ for all markets. Together, these assumptions shut down the sampling error that makes realized within-nest market share ranks differ from expected ranks. For two plans j and k in nest h , $1[r_{jm} > r_{km}] = 1$ without sampling error when $E[1[r_{jm} > r_{km}] | \beta, \vec{x}_m] > \frac{1}{2}$ with sampling error.

Lemma 2 presents a moment inequality. For two plans in the same nest, the plan with the higher mean utility must have the higher market share. An equivalent mathematical operation to finding parameters that satisfy the moment inequalities is finding the set of parameters that maximize the objective function

$$Q_M(\beta) = \frac{1}{M} \sum_{m=1}^M \sum_{h=1}^{H_m} \sum_{j=1}^{J_{hm}} \sum_{k=1, k \neq j}^{J_{hm}} 1[r_{jm} > r_{km}, x'_{jm}\beta - p_{jm} > x'_{km}\beta - p_{km}]. \quad (2)$$

The objective function uses data on plan characteristics and market share ranks for M markets. For each market m , the objective function sums over the H_m pre-specified nests of plans. Each nest h has J_{hm} plans and the estimator compares all pairs of plans j and k . In the data and ignoring ties, one of the two products j or k has a greater market share rank r_{jm} . If in the data $r_{jm} > r_{km}$, for a trial guess of β the estimator asks if indeed j has a higher mean payoff $x'_{jm}\beta - p_{jm}$. If j does, the moment inequality in Lemma 2 is satisfied, and the objective function increases by 1.

As the goal is to maximize the count, or score, of correct predictions of Lemma 2, equation (2) is a maximum score objective function. Without sampling error, the objective function should be maximized with a value equal to its highest theoretical value, which is

$$Q_M^* = \frac{1}{M} \sum_{m=1}^M \sum_{h=1}^{H_m} \frac{J_{hm}(J_{hm} - 1)}{2}.$$

As there are a finite number of inequalities, typically there might be a set of parameter vectors β that maximize the objective function. The identified set B_M^0 of the willingnesses to pay is the set of parameter vectors that maximize the objective function, so that

$$B_M^0 = \arg \max Q_M(\beta).$$

The identified set B_M^0 has no sampling error in it. Identification is not stochastic, so there is no need to prove consistency or conduct hypothesis tests. Either a parameter vector is in B_M^0 and part of the identified set, or it is not.

Consider now that we add data on N new geographic markets, so that we have $M + N$ total markets. As $Q_{M+N}(\beta)$ has to satisfy more inequalities at a maximum, the new identified set of parameter vectors, B_{M+N}^0 , is a subset of the old B_M^0 . For the case of binary choice with microdata, Manski (1988) discusses point versus set identification and derives conditions about the across-market variation in plan characteristics x_m under which sampling new markets allows the identified set B_M^0 to converge to the true parameter in the data generating process, β^0 .

In our Amazon data, the set of plans is identical across markets, with one notable exception. It is likely that each carrier designs a default national set of plans to offer in all markets, and then makes small adjustments based upon regional conditions. Our data do not have the property that sampling new markets gives much variation in the set of plans in a market. Collecting data on different geographic markets at the same point in time is not representative of the asymptotic argument needed for point identification. Even without sampling error in market share ranks, our parameters are only set identified.

It is possible that the objective function is not Q_M^* when evaluated at a maximum. Then the model is misspecified if there is no sampling error in market share ranks. We will now allow for additional errors, so that the model can explain the data when there is sampling error.

3.5 Set Estimation with Sampling Error

There are several reasons why market share ranks can vary across geographic markets with identical plans. For reasons of statistical precision, we do not explore the conjecture that the willingness to pay, β , of consumers varies across markets. We do allow for there to be plans and market specific shocks ξ_{jm} that are common to all customers in a market. Finally, our data come from Amazon, as we will explain. Amazon overweights recent sales in calculating sales ranks. The number of customers purchasing plans on the site, I_m for market m , over these short time horizons is probably not very high. As each customer has idiosyncratic errors ε_{ijm} , a recorded market-share rank may not equal its long-run value.

Both ξ_{jm} and ε_{ijm} induce sampling error in the relative market share ranks of plans within the same nest. This section proves that our estimator is consistent for the identified set $B_M^0 = B^0$ as the number of markets M goes to infinity. For simplicity, we assume that all markets have the same set of J plans with characteristics \vec{x} , and the only differences in the market shares of plans across geographic markets arise from product market specific shocks and because a small number of customers logged into Amazon's site and purchased plans during the time interval Amazon uses to calculate plan market shares. Our lemma in this section focuses only on sampling error and not point identification.

Our maximum score objective function without variation in plan characteristics across markets is

$$Q_M^2(\beta) = \frac{1}{M} \sum_{m=1}^M \sum_{h=1}^{H_m} \sum_{j=1}^{J_{hm}} \sum_{k=1, k \neq j}^{J_{hm}} 1 \left[r_{jm} > r_{km}, x'_j \beta - p_j > x'_k \beta - p_k \right]. \quad (3)$$

Our aggregate data, maximum score set estimator is the set of parameter vectors that maximize $Q_M^2(\beta)$.

We return to the technical details of set estimation and set inference in Section 4. Here we prove a necessary condition for both estimation and inference: the probability limit of our objective function with sampling error is uniquely maximized by the parameters in the identified set. In other words, our model is identified even under sampling error, and in the limit the introduction of sampling error does not alter the identified set.

Lemma 3. *Under Assumption 1, the set of parameters that maximize the probability limit, as $M \rightarrow \infty$, of $Q_M^2(\beta)$ is B^0 , the identified set.*

The proof of identification under sampling error is in an appendix. Note that because of sampling error, the theoretical maximum number of inequalities will not be satisfied, even in the limit. The identification proof works by adding more inequalities from new markets, rather than eliminating sampling error for a fixed number

of markets. Note also that the maximum score objective function is essential for the identification argument. For example, modifying the objective function by putting in a penalty term that weights inequality violations by the amount of the violation would prevent identification.

3.6 Identification Despite Amazon Not Offering All Plans

Amazon offers plans by mostly large national carriers. Smaller carriers are typically present in the offline market. For example, US Cellular sells subscription plans in Chicago that are not offered on Amazon. Even for carriers that do sell on Amazon, not all plans available in the open market are available on Amazon. For example, Cingular offers BlackBerry plans, but T-Mobile is the exclusive BlackBerry retailer on Amazon.

We can identify the willingness to pay parameters β even though Amazon does not offer all plans. Many parametric demand models, such as the multinomial probit and random coefficients logit, are not consistent under similar conditions. Say each consumer i in market m has an idiosyncratic and unknown probability γ_{im}^1 of being aware that Amazon offers mobile phone plans. The unknown probability γ_{im}^1 can be correlated with any observables and unobserved agent-specific nest fixed effects. Further, let there be a probability γ_{im}^2 that, given awareness that the same plan is offered on Amazon and off Amazon, that a consumer buys the plan on Amazon. The probability that consumer i buys plan j in market m is $P_{ijm}^{A_m}$, where the consumer considers A_m , the set of plans offered by all carriers, on and off Amazon. Amazon offers J_m plans in market m . Assume (as is true in our data) that all plans available on Amazon are available from other retailers. The probability that consumer i buys plan j offered on Amazon using Amazon is then $\gamma_{im}^1 \cdot \gamma_{im}^2 \cdot P_{ijm}^{A_m}$. Consumer i 's probability of purchasing plan j using Amazon conditional on i buying one of Amazon's J_m plans using Amazon is

$$P_{ijm}^J = \frac{\gamma_{im}^1 \cdot \gamma_{im}^2 \cdot P_{ijm}^{A_m}}{\sum_{k \in J_m} \gamma_{im}^1 \cdot \gamma_{im}^2 \cdot P_{ikm}^{A_m}}.$$

By canceling terms, it is easy to see that $P_{ijm}^J > P_{ikm}^J$ exactly when $P_{ijm}^{A_m} > P_{ikm}^{A_m}$. Ignoring sampling error, this property that relative choice probabilities and market share ranks are preserved on Amazon means that our estimator will consistently estimate the willingness to pay for plan characteristics.

Fox (2006) introduces and more formally proves this property for the microdata maximum score estimator. Consistency is preserved because maximum score, and by extension our Assumption 1, involves only comparisons between pairs of products in the same nest of choices. Relative choice probabilities and hence market share ranks for products in the same nest are preserved by conditioning on the event that an agent purchased one of the pair of products. By contrast, the multinomial probit and random coefficients logit impose functional form assumptions for the distribution of heterogeneity that do not survive conditioning on an endogenous outcome. Conditioning on an endogenous outcome induces correlation between observables and unobservables, resulting in inconsistency due to selection.

3.7 Comparison to the BLP Assumptions

Berry, Levinsohn and Pakes (1995), or BLP, introduce a set of econometric assumptions and estimation procedures that are now accepted as conventional when working with aggregate data on product characteristics and market share levels. For applications of the BLP framework, see Nevo (2001) and Petrin (2002). It is instructive to consider how the error term properties in Assumption 1 relate to those in BLP.

3.7.1 Market Share Levels vs. Ranks

In BLP, the error term ε_{ijm} is typically assumed to have the type I extreme value or logit distribution, with known parameters. The logit model imposes the scale normalization that the variance of the ε_{ijm} terms is $\pi^2/6$. We make the alternative scale normalization that the coefficient on price is -1 , as we cannot estimate the distribution of the ε_{ijm} with data on market share ranks, as we now explain.

With the logit distributional assumption and a large number of customers in each market, market shares have a known functional form. Picking some alternative l as a reference allows the estimation of the price and non-price coefficients, α and β , using the regression equation

$$\log s_{jm} - \log s_{lm} = (x_{jm} - x_{lm})' \beta - \alpha(p_{jm} - p_{lm}) + (\xi_{jm} - \xi_{lm}). \quad (4)$$

Note how a small price change Δp_{jm} decreases $\log s_{jm} - \log s_{lm}$ by a specific amount: $-\alpha$. By contrast, a small price change Δp_{jm} may leave market share ranks r_{jm} unchanged. Thus the assumption of a known distribution for the error term allows one to more easily identify the parameters of interest. Data on market share ranks cannot make use of this shape information, to the point that the finite sample objective function generically has a set of global maximizers. The requirements of market share rank data further highlight how parametric distributional assumptions give point identification and lower standard errors in traditional applications.

3.7.2 Correlation of Price with Unobserved Product Attributes

BLP use instruments to control for the price endogeneity from the correlation of p_{jm} with the unobserved ξ_{jm} . If instruments are available, one can control for price endogeneity using a maximum score estimator as well. Fox (2006) introduces and proves the consistency of a maximum score instrumental variables estimator.

We do not have instruments for price (such as carrier-specific cost shifters) in our mobile phone data from Amazon. We do allow for consumer and nest fixed effects v_{ihm} . For mobile phone calling plans, v_{ihm} captures features that vary only by carrier, such as free in-network calling on Verizon. Because the fixed effects are agent specific, one customer can value carriers differently than another. In mobile phones, a vertical unobservable ξ_{jm} alone is inadequate because one carrier's coverage may be better in one neighborhood and worse in another.

Once consumer and carrier fixed effects are controlled for, there are no unobserved product characteristics in mobile phone calling plans. The difference between two calling plans from the same carrier is only in the observable characteristics we include in our estimates. A plan is a collection of legal details; there is no workmanship in a plan causing an unquantifiable quality that may be correlated with price.

3.7.3 Heterogeneity Over Tastes for Product Characteristics

In many applications, consumers have heterogeneous tastes over the observable characteristics of products. For mobile phone calling plans, consumers are very likely to have heterogeneous willingnesses to pay to talk on the phone. A salesperson with a lot of clients may be willing to buy an expensive plan that offers 6000 minutes of daytime calling a month, while a person who uses his or her phone only in emergencies may prefer the plan with the least amount of minutes.

BLP would capture heterogeneity in the willingness to pay for anytime minutes using a random coefficient specification. No scholar has generalized results for the two-choice model from Manski (1975) about the

semiparametric identification of the mean willingness to pay under random coefficients to the case with more than two choices, which was also studied by Manski (1975). Given this, researchers working with market share rank data cannot use a random coefficients specification.

We use our nest fixed effects to control for the heterogeneous willingness to pay for anytime minutes. For each carrier, we compare a regional plan only to the national plans with slightly more and slightly fewer minutes. For example, Verizon offers national plans with 450, 900, 1350, 2000, 4000 and 6000 anytime minutes a month. Verizon also offers regional plans with 600, 1200 and 1800 minutes. We use the following comparisons: 450 to 600, 600 to 900, 900 to 1200, 1200 to 1350, 1350 to 1800, and 1800 to 2000. We order these seven plans, so that $J_{hm} = 7$ for Verizon. Our maximum score objective function is

$$Q_M^3(\beta) = \frac{1}{M} \sum_{m=1}^M \sum_{h=1}^{H_m} \sum_{j=1}^{J_{hm}-1} \left(1 \left[r_{jm} > r_{j+1,m}, x'_j \beta - p_j > x'_{j+1} \beta - p_{j+1} \right] + 1 \left[r_{jm} < r_{j+1,m}, x'_j \beta - p_j < x'_{j+1} \beta - p_{j+1} \right] \right). \quad (5)$$

We difference out heterogeneity in the demand for anytime minutes to the greatest extent possible.

4 Set Inference

Our objective function is set identified as M goes to infinity. There are several recently-developed methods for performing inference on set-identified estimators, for example Andrews, Berry and Jia (2005), Beresteanu and Molinari (2006), Imbens and Manski (2003), Pakes, Porter, Ho and Ishii (2005), Rosen (2006), and Shaikh (2005). We will follow the methods developed in Chernozhukov, Hong and Tamer (2005), or CHT.

A contour set of the finite-sample objective function, $C_M(c)$, is defined as all parameter vectors that generate objective function values above a threshold, c ,

$$C_M(c) = \left\{ \beta : Q_M^3(\beta) - \sup_{\beta'} Q_M^3(\beta') \geq c/a_M \right\}. \quad (6)$$

Here we normalize the objective function by the best fit, as the limit of the objective function is not maximized at Q^* , as was the case in the model without sampling error. By subtracting the finite-sample maximum, our objective function achieves its maximum at 0, as in CHT. Consequently, all relevant thresholds c will be nonpositive. Kim and Pollard (1990) have demonstrated that $a_M = M^{2/3}$ is an appropriate normalization for estimators such as the maximum score.

For some confidence level α , CHT's procedure finds a confidence contour set $C_M(c_\alpha)$ for the identified set, B^0 , such that

$$\liminf_{M \rightarrow \infty} P(B^0 \subseteq C_M(c_\alpha)) \geq \alpha.$$

The set of parameters $C_M(c_\alpha)$ is a 95% confidence set when we substitute an appropriately chosen critical threshold c_{95} . We now turn to computing c_α .

CHT propose two alternative collections of conditions under which their procedure is valid. The first set of conditions requires that the probability limit of the objective function be continuous in the unknown willingness to pay parameters β . One can examine the probability limit of our objective function in the proof of Lemma 3. As β only enters the probability limit of the objective function in a finite number of step functions, the

probability limit is not continuous in β .

CHT’s alternative regularity condition involves estimators that have degenerate interior asymptotics: the function has an asymptotic distribution that vanishes over the interior of the identified set. Degenerate asymptotics often occur in moment inequality methods that are related to but are not identical to maximum score. The degenerate interior condition is satisfied for our objective function. Under the degenerate interior property, we can form a consistent estimator of the identified set (not a 95% confidence interval) by substituting $c = 0$ into equation (6).

Computationally, we cannot search over a continuous space. We first draw a large number of candidate parameter vectors in the neighborhood of B_M . We modify CHT’s suggestion of using the Laplace-type estimators (LTE) proposed in Chernozhukov and Hong (2003). The LTE is a Markov Chain Monte Carlo (MCMC) technique useful for both optimizing and performing inference on numerically-difficult statistical functions. Instead of generating a Markov chain with the statistical properties as outlined in Chernozhukov and Hong, we instead accept all potential parameter vectors that generate objective function values equal to $c = 0$. We can choose the cutoff of 0 because of the degenerate interior asymptotics, as discussed in CHT. By exploiting the degenerate interior asymptotics, we know that the Markov chain gives us a large set of parameter vectors that will cover the identified set in the limit. The MCMC procedure gives us an estimator of the identified set that we call $C_M(0)$.

We now turn to computing c_α for inference. CHT suggest a statistical resampling procedure known as subsampling. We generate a large number of subsamples drawn from the $M = 22$ markets in our data set. In each subsample, $M_{\text{sub}} = 5$ markets are drawn without replacement. We repeat and generate 1000 subsamples, indexed by b . We take the 1000 subsamples and generate the following set of 1000 normalized objective function values:

$$\left\{ \inf_{\beta \in C_M(0)} \left(a_{M_{\text{sub}}} Q_b(\beta) - \sup_{\beta' \in C_M(0)} a_{M_{\text{sub}}} Q_b(\beta') \right) \right\}_{b=1}^{1000}.$$

In other words, we take all of our parameter vectors generated by the modified LTE process and collect the worst fit in our candidate set $C_M(0)$ for each of the 1000 subsamples. As is done in subsampling, we take the 5th percentile of these objective function fits as our estimator of c_α . Our 95% estimator is then the level set $C_M(c_\alpha)$ from equation (6). In tables, we describe our 95% confidence set $C_M(c_\alpha)$ by reporting the projection of the extremal elements of the set onto the real line.

5 Amazon Data

5.1 Markets and Plans

Data availability has been the most challenging impediment to studying demand-side antitrust issues in the wireless phone industry. Carriers release annual, nationwide data on total subscribers, not market-level subscribers. Also, carriers do not report on the popularity of their individual subscription plans. The FCC has access to the sets of phone numbers given to carriers, and the FCC uses the information to approximate market shares. However, the FCC does not release its confidential data to researchers. We asked.³

³Some companies collect phone bills from consumers. Bill harvest data are not entirely appropriate, as at any given point in time the stock of mobile phone users has plans purchased from the menus of plans available in many different time periods. The time of plan

The online retailer Amazon sells mobile phones and attached subscription plans. An online retailer faces a disadvantage because consumers cannot physically examine a phone as they can in a brick-and-mortar store. Amazon's competitive advantage is that it offers deep mail-in rebates on phones, so its prices for buying a new phone (when attached to a new one or two year subscription) are often lower than the prices on a carrier's own website. Amazon does not discount the monthly fees of subscription plans. Monthly fees are billed directly by the carrier providing service.

We collect detailed plan characteristic and market share rank data from Amazon's site. For plan j in market m , we observe the plan's monthly fee, p_{jm} , and a vector of d other plan features x_{jm} . Many features, such as the one-time activation fee, are constant across the comparisons that we include in our maximum score objective function. Amazon rank orders the top selling plans for different geographic markets, so we observe the rank order of the market share of each Amazon plan in each geographic market. The rank order of plan j in market m is r_{jm} .⁴

A carrier must own a FCC license in a geographic market to legally enroll subscribers there. The competitors in each market may differ, as a fixed number of licenses are issued per market. Therefore, the competition for customers is primarily local. The set of plans offered by carriers is relatively constant across markets; the choice of the menu of plans appears to primarily operate on the national level. We collect data on different markets to increase statistical precision.

We collected data for 22 of the largest metropolitan areas in the United States.⁵ The boundaries of markets are chosen by Amazon and we do not know them. However, the boundaries appear to correspond to the popular notion of a metropolitan area. For example, the same plans and market share ranks appear for the nearby cities of Los Angeles and Riverside, CA, but the plans and market share ranks differ for San Diego, CA, which is typically considered a separate metropolitan area. We use only markets that both Amazon and we agree are separate metropolitan areas. We have verified that the market share ranks are not volatile over a week, particularly for the plans with the highest sales ranks. Ghose and Sundararajan (2006) mention that Amazon uses a rolling window of sales to compute market share ranks. In the summer of 2004, Amazon moved to a system that calculates market share ranks using exponential decays that give more weight to newer purchases.

Amazon carries plans from all five national carriers: Cingular, Nextel, Sprint, T-Mobile and Verizon. The smaller carrier EarthLink Wireless offers plans on Amazon in 15 of the 22 markets.⁶

Table 1 lists the number and characteristics of the plans that we use in estimation. Only carriers with some regional plans provide variation that can identify the willingness to pay for national coverage. Table 1 shows that a typical market has seven T-Mobile plans, six or nine Verizon plans, and six or nine Verizon family plans.

Table 1 also describes the number of plans that have only regional coverage. A plan can allow a traveling customer to make calls from across the United States with no surcharges (national coverage), or a carrier may purchase may not be observable.

⁴If a researcher has another dataset with continuously measured shares s_{jm} , it is easy to convert those shares into ranks.

⁵We have used the Amazon site extensively and wish to explain a little of how the site worked in late 2005, when the data were collected. When you go to the site to shop for mobile phone plans, you are prompted to enter your zip code. Amazon uses the zip code to look up your geographic market. We collect market share rank data by choosing a zip code corresponding to each city. There are various pages. The page that presents the plans rank ordered by sales is reached by using the toolbar to search "Wireless Plans" for a blank string. All plans will appear, and you can sort them by sales rank. The resulting plans are presented in a matrix, with the top three plans in the first row, plans three to six in the second row, etc. A plan's rank comes from its position in the matrix, not from a text label, as Amazon includes for books. You can verify the ordering of plans in the matrix by consulting another page, which lists the top five plans in numeric order, with the rankings listed explicitly. We clicked on each plan and manually copied its characteristics.

⁶Amazon sells prepaid service, where a customer does not pay a set monthly fee. We do not consider the data on prepaid service, because plan characteristics such as price and anytime minutes are not comparable to the monthly values for subscription plans. A customer that uses all of his or her monthly minutes will find it cheaper to subscribe to a monthly plan.

Table 1: Plans in a Market Used in Estimation

Carrier Nest	# of Plans	# of Markets	Minute Range	Price Range	# Regional
T-Mobile National	7	22	300—5000	\$29.99—\$129.99	1
Verizon National	6	22	450—6000	\$39.99—\$199.99	0
Verizon Regional	3	7	600—1800	\$49.99—\$89.99	3
Verizon Family National	6	22	700—6000	\$60—\$290	0
Verizon Family Regional	3	7	600—1800	\$49.99—\$89.99	3

levy such surcharges on travelers (regional coverage). Most plans in our data offer national coverage. However, there are important regional plans. T-Mobile offers one regional plan with 3000 minutes. T-Mobile charges a regional subscriber an extra 49 cents a minute when using his or her phone outside his or her home region. Verizon offers three regional individual plans and three regional family plans. Both family and individual plans charge 69 cents a minute for roaming. The Verizon regional plans are offered only in markets in the West: Denver, Los Angeles, Phoenix, Portland, San Diego, San Francisco and Seattle. The Verizon regional plans are only present in the West offline as well; Verizon does not offer regional plans in the remainder of the country.

All regional plans levy long distance surcharges when a customer phones a customer in another market. T-Mobile and Verizon both charge 20 cents a minute for long distance. Verizon allows regional subscribers to upgrade to free long distance for \$5 a month.⁷

We will estimate the willingness to pay for national coverage, the opposite of regional coverage. We only compare plans to plans in the same carrier nest and in the same market. We also only compare a plan to plans with slightly more or fewer anytime minutes. Therefore, our estimator uses only within-market variation. The fact that Verizon regional plans are not found in all markets is not a source of variation that our maximum score estimator exploits.

5.2 Popularity of National Coverage

Each plan in our sample has a market share rank r_{jm} . We change market share rank from an integer to a percentile, by the formula

$$\tilde{r}_{jm} = \frac{r_{jm}}{\max_{j \in J_m}(r_{jm})}.$$

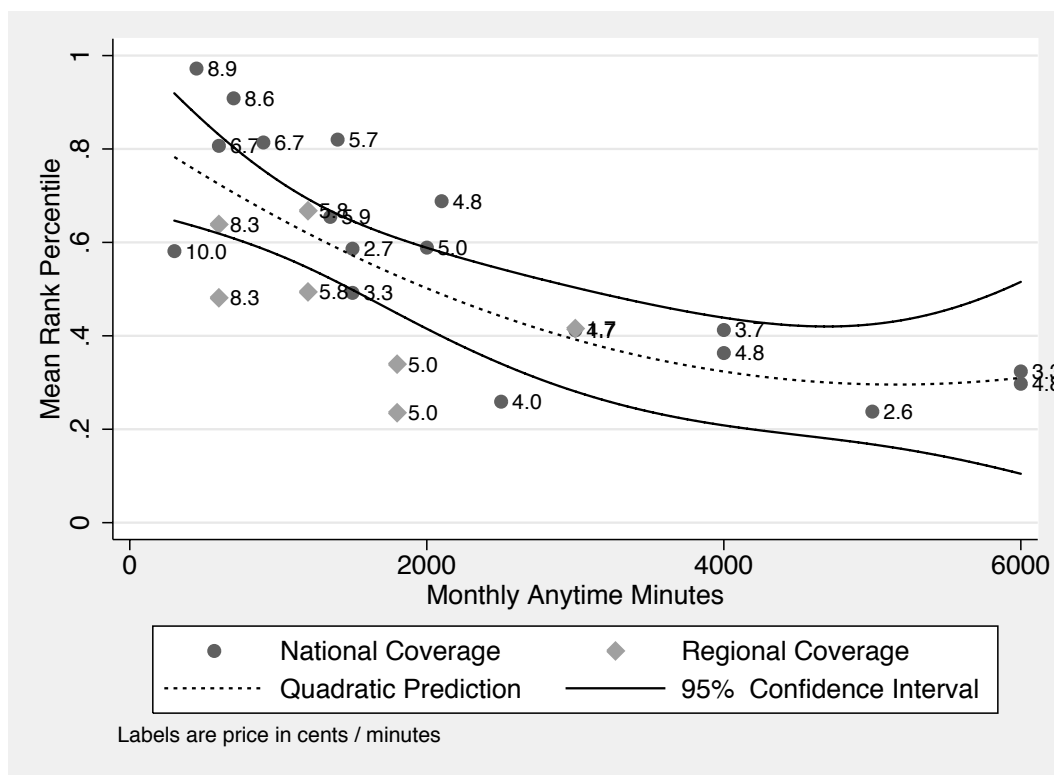
This normalizes the ranks, so that the most popular plan in a market has a percentile rank of 1, and the least popular plan has a percentile rank of close to 0. We then compute the mean of each plan’s percentile rank across the markets where the plan is offered. Most plans are offered in all markets.

Table 1 shows T-Mobile, Verizon and Verizon Family offer plans with regional coverage. For those three carrier nests, Figure 1 plots the mean percentile ranks of each plan by its monthly anytime minutes, along with a fitted quadratic and its confidence interval. The plan labels are monthly prices in cents divided by monthly anytime minutes. The figure compares plans across the three carrier nests, which we will not do in our willingness to pay estimates.

First note that plans with more monthly minutes are less popular. Plans with many minutes cost less per minute: the 6000 minute Verizon individual plan costs 3.3 cents per minute, while the 450 minute plan costs

⁷While not in the table and our estimation sample, Nextel offers four plans that do not include free long distance. Nextel uses a proprietary phone technology that prohibits its customers from operating on networks owned by almost all other carriers. Consequently, Nextel does not levy charges to travelers, in part because its phones are incapable of operating off its network. We do not consider Nextel. Also, Cingular dropped its regional plans from Amazon just before data collection began.

Figure 1: Popularity of National and Regional Plans



8.9 cents per minute. High minute plans are unpopular not because they are a bad deal, but because many consumers do not have such a strong need to talk on the phone. To address this unobserved heterogeneity in demand for anytime minutes, our structural estimator uses only comparisons between plans with similar numbers of minutes, as discussed in Section 3.7.3.

Figure 1 shows that plans with regional coverage are less popular than plans with national coverage. Four of the seven regional plans are below the fitted quadratic's 95% confidence interval, and three plans are within the 95% confidence interval. No regional plans lie above the confidence interval, although many national plans do.

Verizon individual regional plans are always less popular than similar national plans. For example, the 600 minute Verizon regional plan has a mean percentile rank of 0.64, which is much lower than the ranks of 0.97 for the 450 minute national plan and 0.81 for the 900 minute national plan. On a price per minute basis, the Verizon regional plans are bad deals. The 450 national plan charges 8.9 cents per minute, while the 600 minute regional plan charges 8.3 cents per minute.

T-Mobile discounts regional plans more. The 3000 minute T-Mobile regional plan charges 1.7 cents per minute, while the 2500 and 5000 minute plans with national coverage charge 4.0 and 2.6 cents per minute, respectively. The 3000 minute regional plan is more popular, with a mean percentile rank of 0.42 compared to the ranks of 0.26 and 0.24 respectively for the 2500 and 5000 minute national plans. Nonetheless, the low price per minute of the T-Mobile regional plan suggests that consumers may have a positive willingness to pay

for national coverage.

5.3 Representativeness of Amazon Users

Customers who use Amazon differ in preferences from non-users. Section 3.6 shows that our estimator is consistent for the population willingness to pay even if people have different probabilities of knowing about Amazon’s cell phone offerings, and if plans from some carriers are not offered on Amazon.

Another possibility is that Amazon users have a higher willingness to pay for national coverage. To address this, we turn to an auxiliary dataset on internet use. The market-research firm Forrester surveyed 68,664 Americans in its 2005 Technographics Benchmark survey. The Forrester data oversample heads of household, as only 2.3% of reported mobile phone users in the survey are under 25. As Amazon has no salespeople to field questions from new users, we suspect most mobile phone customers on Amazon are upgrading to a new phone, not buying a phone for the first time. We use the roughly 70% of respondents that report owning a mobile telephone as our base sample. Of mobile phone users, 40% have purchased an item online at least once in the past twelve months, and 13% have “shopped” at Amazon itself in the last 30 days. Unfortunately, we cannot isolate the presumably small sample of people who purchased phones on Amazon.

Among all mobile phone users, Amazon shoppers are younger, wealthier and actually not much more likely to be male. The Forrester data suggest that 31% of Amazon phone users are under 40, while only 21% of non-Amazon users are under 40. Using a midpoint approximation to a survey question, Amazon households earn \$23,000 more in a year than non-Amazon households. One stereotype that is not true is that Amazon shoppers are not much more likely to be male: 48% of Amazon shoppers are male, compared with 46% of non-Amazon mobile phone users. 41% of Amazon users have children under 18 at home, compared with 36% of non-users.

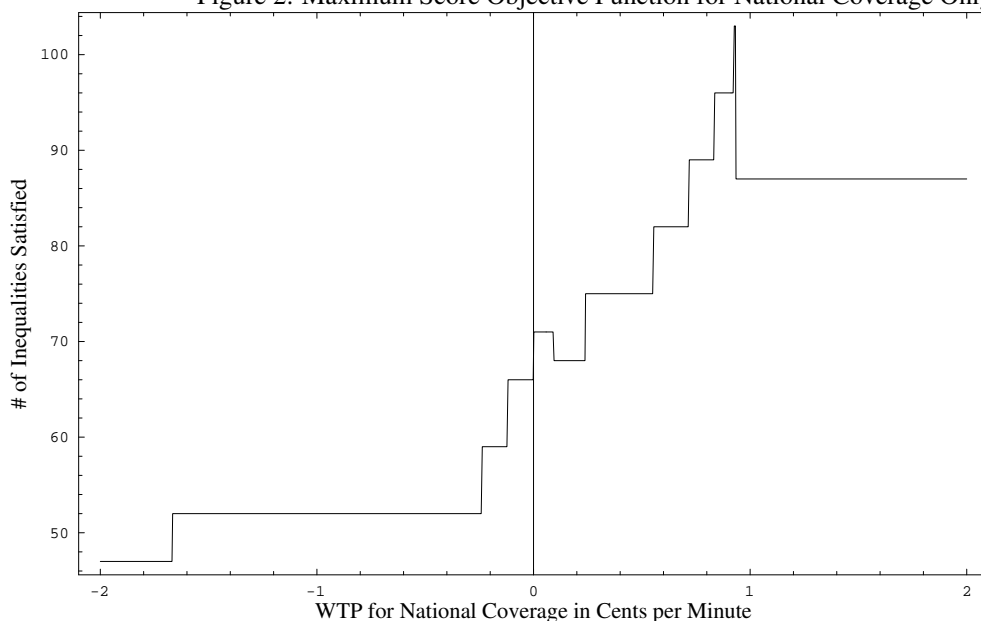
We estimate the willingness to pay for national coverage. National coverage is only valued by travelers. Amazon users are more likely to travel than non-users. 33% of Amazon users report going on one or more business trips in the past twelve months, compared to 18% of non-users. Similarly, 47% of Amazon users have recently gone on a pleasure trip, compared to 33% of non-users. It is likely that our estimate of the valuation of national coverage is an upper bound, given that Amazon users are more likely to travel than non-users.

Amazon is a useful price comparison tool, so customers on Amazon may be savvier. One piece of evidence suggesting the customers using Amazon are savvier is that Amazon often offers more aggressive rebates on phones than the carriers themselves. One can save money on phones (but not plans) by shopping on Amazon.

6 Willingness to Pay

We estimate the willingness to pay for national coverage and other plan characteristics by optimizing the maximum score objective function, equation (5). Equation (5) compares a plan only to plans with slightly more and fewer minutes from the same carrier nest in the same geographic market. This procedure controls for heterogeneous preferences for phones, for coverage near someone’s residence and workplace, and for anytime minutes.

Figure 2: Maximum Score Objective Function for National Coverage Only



6.1 National Coverage Without Other Controls

This section explores a specification where national coverage is the only characteristic other than the monthly fee divided by the number of anytime minutes that enters the objective function.

Figure 1 shows the plot of the maximum score objective function when we include only one non-price characteristic: national coverage. The objective function attains its maximum of 103 out of 120 inequalities at willingness to pay for national coverage from between 0.926 and 0.933 cents per minute. The mean price per minute of the 25 plans from T-Mobile, Verizon and Verizon Family is 5.44 cents per minute. So the willingness to pay for national coverage is 17% of the monthly price per minute of subscription plans.

We will discuss the implications of the willingness to pay (WTP) for national coverage for evaluating mobile phone mergers below. For now, we wish to explain the characteristics and relative popularities of plans that drive the estimate of a WTP to pay national coverage of 0.926–0.933 cents per minute.

Table 1 shows that there are seven subscription plans that offer regional coverage. T-Mobile offers one plan and Verizon and Verizon Family each offer three. As our estimator compares only plans from one carrier to plans with the next highest and lowest numbers of minutes, the identification of the parameter for national coverage arises only from variation across plans within the carrier nests. We consider each carrier nest separately.

T-Mobile offers one regional plan in all 22 markets. The regional plan offers 3000 minutes for \$49.99, which comes to 1.67 cents a minute. Our estimator compares this 3000 minute regional plan to a 2500 minute national plan that costs 4.00 cents a minute, and a 5000 minute national plan that costs 2.60 cents per minute. The across-market popularity in decreasing order is the 3000, the 2500 and finally the 5000 minute plan. Given the realized within-market popularities, and ignoring division by the number of markets, our objective function for T-Mobile is

$$19 \cdot 1 [\beta^{\text{Nat}} < 0.933] + 21 \cdot 1 [\beta^{\text{Nat}} < 2.333] + 1 [\beta^{\text{Nat}} > 2.333] + 3 \cdot 1 [\beta^{\text{Nat}} > 0.933].$$

The number 0.933 is the difference in price per minute between the 3000 and 5000 minute plans, and the number 2.333 is the difference in price per minute between the 2500 and 3000 minute plans. In 19 out of 22 markets, the 3000 minute plan is more popular than the 5000 minute national plan, so national coverage must be worth less than 0.933. Any value of national coverage less than 0.933 maximizes the T-Mobile objective function at 40 out of 44 inequalities. One objection might be that 5000 minutes is a lot more than 3000 minutes. A T-Mobile subscriber with a strong need for around 5000 minutes might never consider the 3000 minute plan. In this case, the only valid comparison is with the 2500 minute plan. The regional plan is more popular in 21 out of 22 markets. We can then upper bound the willingness to pay for national coverage at 2.33 cents.

Verizon offers three non-family regional plans and six non-family national plans. The regional plans are offered only in seven Western cities. The objective function for Verizon is

$$1 [\beta^{\text{Nat}} < -1.67] + 6 \cdot 1 [\beta^{\text{Nat}} > -1.67] + 1 [\beta^{\text{Nat}} < 0.0] + 6 \cdot 1 [\beta^{\text{Nat}} > 0.0] + 5 \cdot 1 [\beta^{\text{Nat}} < 0.0927] + 2 \cdot 1 [\beta^{\text{Nat}} > 0.0927] + 7 \cdot 1 [\beta^{\text{Nat}} > 0.926] + 7 \cdot 1 [\beta^{\text{Nat}} > 0.833] + 7 \cdot 1 [\beta^{\text{Nat}} > 0.555]. \quad (7)$$

The Verizon objective function is maximized by any β^{Nat} greater than 0.926 cents per minute. The number 0.926 is the difference in the cents per minute of the 1350 minute national plan (5.925) and the 1800 minute regional plan (4.999). The 1800 minute regional plan is less popular in six out of seven markets than the 1350 minute national plan, and less expensive per minute, so we can form only a lower bound on β^{Nat} . However, a reasonable person might suspect the reason the 1800 minute regional plan is not very popular is a 2000 minute national plan with the same price per minute: 4.999. Verizon regional plans tend to have one very near neighbor national plan in terms of anytime minutes. If we only compare each Verizon regional plan to its single nearest neighboring national plan, the objective function for the three regional plan comparisons becomes

$$7 \cdot 1 [\beta^{\text{Nat}} > 0.555] + 5 \cdot 1 [\beta^{\text{Nat}} < 0.0927] + 2 \cdot 1 [\beta^{\text{Nat}} > 0.0927] + 1 [\beta^{\text{Nat}} < 0.0] + 6 \cdot 1 [\beta^{\text{Nat}} > 0.0]. \quad (8)$$

By narrowing our set of comparison national plans, our lower bound for β^{Nat} becomes 0.555, which is the price per minute difference between Verizon's very popular 450 minute national plan (8.887 cents) and the somewhat less popular 600 minute regional plan (8.332 cents). In this case, using a more conservative choice of inequalities makes our bounds wider.

Verizon also offers three regional and six national family plans. The objective function is

$$7 \cdot 1 [\beta^{\text{Nat}} > 2.739] + 7 \cdot 1 [\beta^{\text{Nat}} > 0.715] + 7 \cdot 1 [\beta^{\text{Nat}} > 0.240] + 7 \cdot 1 [\beta^{\text{Nat}} > -0.118] + 7 \cdot 1 [\beta^{\text{Nat}} > -0.238].$$

In all cases the regional family plan is less popular than the national family plan. The lower bound for β^{Nat} of 2.739 comes from the largest price difference between two adjacent plans: a 700 minute national plan (8.571 cents) and a 1200 minute regional plan (5.832 cents). A lower bound of 2.739 is inconsistent with the upper bounds of, depending on the comparisons used, of 0.933 and 2.33 from the T-Mobile data. In the pooled sample of all carriers in Figure 2, the T-Mobile plans dominate the Verizon Family plans as the T-Mobile regional plan is sold in all 22 markets, versus 7 for the Verizon Family regional plans.

However, we might want to restrict comparing the 1200 minute regional plan to its nearer 1400 minute

national plan neighbor. By making only the closest possible plan comparisons, our objective function becomes

$$7 \cdot 1 [\beta^{\text{Nat}} > 0.240] + 7 \cdot 1 [\beta^{\text{Nat}} > -0.118] + 7 \cdot 1 [\beta^{\text{Nat}} > -0.238].$$

By using only the most reasonable comparisons, our lower bound for β^{Nat} becomes 0.240, which is the price difference between the 600 minute regional plan (8.332 cents) and the 700 minute national plan (8.571 cents). The number 0.240 is not the tightest bound, as the lower bound from the Verizon individual plans is 0.555.

To conclude, we observe a very popular T-Mobile regional plan. This places an upper bound on the willingness to pay per minute for national coverage. We also observe unpopular Verizon regional plans, which place lower bounds on the willingness to pay for national coverage. We first consider a specification where we compare each regional plan to the national plan with the next fewer anytime minutes and the national plan with the next most anytime minutes. Using the fact that the T-Mobile plans appear in all 22 markets to weight them more than the Verizon Family plans, our bounds are very tight: the willingness to pay for national coverage is between 0.926 and 0.933 cents per minute. If we include only the more conservative comparison of each regional plan to its single most similar national plan neighbor, the willingness to pay for national coverage is bounded between 0.555 and 2.333 cents per minute.

6.2 Controls for Other Regional Plan Features

There are two omitted characteristics of the regional plans considered above. First, the 3000 minute T-Mobile regional plan does not offer free calling during nights and weekends, a popular feature for non-business users. The 2500 and 5000 minute T-Mobile national plans that we compare the regional plan to do have unlimited nights and weekends. Second, both the T-Mobile and Verizon regional plans charge customers for making long distance calls. National coverage involves surcharges for placing calls when a user travels; long distance charges are incurred when a user in any region places a call to a number in a distant region. It is not clear why mergers would affect the ability of carriers to offer free long distance. This section controls for these other characteristics.

T-Mobile offers a 1500 minute national plan for \$39.99 that does not offer unlimited calling during nights and weekends. This compares closely to a 1500 minute national plan for \$49.99 that does offer unlimited nights and weekends. In 13 of our 22 markets (and 12 out of 3 markets where these plans were very popular), the \$39.99 plan is more popular, meaning an estimate for the willingness to pay for unlimited nights and weekends is any value less than \$10 a month. Unlimited nights and weekends adds extra minutes and so is a substitute rather than a complement for anytime minutes. Therefore, the willingness to pay for unlimited nights and weekends is not a structural constant in terms of cents per minute.

If we apply the $\beta^{\text{NightWeek}} \leq \10 willingness to pay to T-Mobile's 3000 minute regional plan, the value per minute for that plan is $0 \leq \beta^{\text{NightWeek}}/30 < 0.333$ cents per minute, where 30 is 3000 minutes per month divided by 100 cents per dollar. Considering that the T-Mobile national plans we compared the regional plan to had unlimited nights and weekends, our previous upper bound becomes $\beta^{\text{Nat}} + \beta^{\text{NightWeek}}/30 < 0.933$. The upper bound on $\beta^{\text{NightWeek}}/30$ does not lower the marginal upper bound on β^{Nat} , which is still 0.933. The theoretical lower bound of 0 for $\beta^{\text{NightWeek}}/30$ does not lower the upper bound on β^{Nat} , either. After examining unlimited nights and weekends, our upper bound is still $\beta^{\text{Nat}} < 0.933$. Likewise, our conservative upper bound remains 2.33. Note that our estimate of $\beta^{\text{NightWeek}} < 0.333 \cdot 30 = \10 a month is an univariate analysis from the two

1500 minute plans. Our full multivariate analysis below tightens this upper bound for $\beta^{\text{NightWeek}}$.

T-Mobile and Verizon's regional plans charge for long distance, in addition to deducting the call length from the standard allotment of minutes. However, Verizon does allow customers to add free domestic long distance to any regional plan for an extra \$5. Theoretically, free long distance should raise the value of an anytime minute by β^{FreeLong} . For whatever reason, Verizon charges a fixed fee. To more accurately find the lower bound on national coverage, we will assume that all Verizon users pay the extra \$5 a month fee. For example, a user who buys the 1200 minute regional plan faces an extra charge of $500/1200 = 0.417$ cents per minute.

Verizon's maximum score objective function when we compare each regional plan to the national plans with the next fewer and next most number of minutes, equation (7), is

$$1 [\beta^{\text{Nat}} < -2.50] + 6 \cdot 1 [\beta^{\text{Nat}} > -2.50] + 1 [\beta^{\text{Nat}} < -0.278] + 6 \cdot 1 [\beta^{\text{Nat}} > 0.278] + 5 \cdot 1 [\beta^{\text{Nat}} < -0.324] + 2 \cdot 1 [\beta^{\text{Nat}} > -0.324] + 7 \cdot 1 [\beta^{\text{Nat}} > 0.648] + 7 \cdot 1 [\beta^{\text{Nat}} > 0.416] + 7 \cdot 1 [\beta^{\text{Nat}} > -0.278].$$

This objective function is maximized by $\beta^{\text{Nat}} > 0.648$. Verizon's conservative maximum score objective function in equation (8) becomes

$$7 \cdot 1 [\beta^{\text{Nat}} > -0.2783] + 5 \cdot 1 [\beta^{\text{Nat}} < -0.324] + 2 \cdot 1 [\beta^{\text{Nat}} > -0.324] + 1 [\beta^{\text{Nat}} < -0.277] + 6 \cdot 1 [\beta^{\text{Nat}} > -0.277].$$

Unfortunately, maximizing this objective function says that $\beta^{\text{Nat}} > -0.277$. If we are conservative in our choice of inequalities, our plan characteristics are not rich enough to distinguish between free long distance and national coverage. National coverage is a positive characteristic, so we set β^{Nat} 's lower bound to 0. The conservative lower bound is uninformative because Verizon's regional plans are an even worse financial deal if one considers the extra \$5 fee for free long distance.

T-Mobile does not allow regional plan users to upgrade to free long distance. For lack of a better solution, we assign Verizon's \$5 fee to T-Mobile 3000 minute regional plan. The per-minute cost is 0.167 cents per minute. Therefore, our tight upper bound for β^{Nat} decreases from 0.933 to 0.766 and our conservative upper bound increases from 2.333 to 2.166. Applying the \$5 fee to the T-Mobile regional plan is a conservative approach: if a consumer with the 3000 minute regional plan valued free long distance more than \$5 a month, the upper bounds would decrease by more and become tighter. In conclusion, our tight bound for β^{Nat} is 0.648 to 0.766 cents per minute. Our conservative bound for β^{Nat} is 0 to 2.166 cents per minute.

Figure 3 shows the final maximum score objective function for pairwise comparisons, evaluated at the \$5 Verizon fee and the theoretical lower bound $\beta^{\text{NightWeek}} = 0$. Our tight bound for β^{Nat} is 0.648 to 0.766 cents per minute is clearly visible in the picture.

6.3 Subsampled Confidence Intervals

Section 4 discusses how we use the Chernozhukov, Hong and Tamer (2005) procedure to estimate 95% confidence regions for each parameter. We subsample by drawing fake datasets composed of subsets of 5 of our 22 markets. This statistical procedure allows for market-specific taste shocks ξ_{jm} and allows for sampling from a small number I_m of customers purchasing calling plans on Amazon. To the extent that market share ranks are similar across markets, our confidence sets will approximate the estimated bounds above.

Figure 3: Maximum Score Objective Function for National Coverage Evaluated at \$5 Long Distance Fee and $\beta^{\text{NightWeek}} = 0$

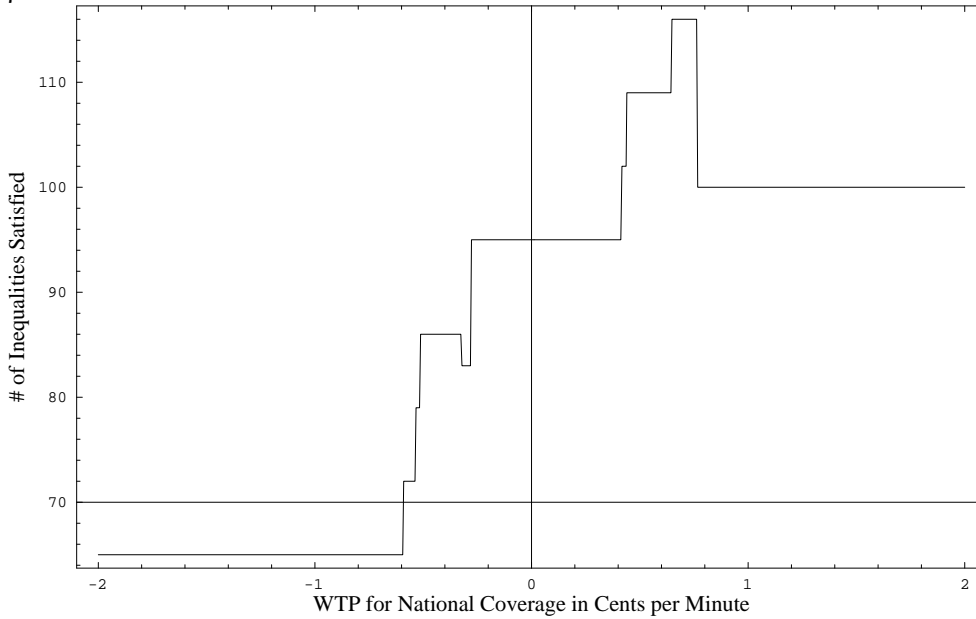


Table 2: Set Estimates and Confidence Regions for WTPs for Univariate and Multivariate Specifications

Variable	Lower	Upper	Length
Set Estimates			
WTP for National (cents/minute)	0.648	0.767	0.119
WTP for National (cents/minute)	0.648	0.767	0.119
WTP for Unlimited Off Peak (\$/month)	0	3.57	3.57
95% Confidence Region			
WTP for National (cents/minute)	0.648	0.767	0.119
WTP for National (cents/minute)	0.648	0.767	0.119
WTP for Unlimited Off Peak (\$/month)	0	3.57	3.57

We incorporate Section 6.1’s comparisons of each regional plan to two national plans and Section 6.2’s comparison of the 1500 minute T-Mobile plan without unlimited night and weekends to the 1500 minute plan with unlimited nights and weekends. As before, β^{Nat} enters the value of an anytime minute, while $\beta^{\text{NightWeek}}$ is a monthly value, as unlimited nights and weekends are a substitute for anytime minutes. We apply the \$5 Verizon fee for free long distance calling to all regional plans.

Table 2 reports set estimates as well as 95% confidence regions from the subsampling approach discussed in Section 4. For each method, we report two specifications: one with only national coverage and one controlling for unlimited off peak calling. We see that the 95% confidence region for the willingness to pay for national coverage has the same tight bounds as reported before: 0.648 to 0.767 cents per minute. The 95% confidence region is the same as the set estimate because of the strong similarity in market share ranks across our sample of 22 markets: the estimates using a subset of the markets are typically the same as for the entire sample.

The comparison between the 1500 minute T-Mobile plans is not informative for the lower bound for the

willingness to pay for unlimited off peak calling, so we set the lower bound's value to the theoretical constraint of 0 in the subsampling. Previously, our univariate analysis of the two 1500 minute plans yielded an upper bound for the willingness to pay of unlimited off peak calling of $\beta^{\text{FreeNight}} = \10 a month, the difference in price between the two plans. Our specification with both the WTPs for national coverage and unlimited nights and weekends is multivariate. The value of unlimited nights and weekends enters the objective function for the T-Mobile 3000 minute regional plan as well. The upper bound on the value of unlimited nights and weekends comes from evaluating the maximum score objective function at the lower bound for the WTP for national coverage of $\beta^{\text{Nat}} = 0.648$. The objective function is

$$19 \cdot 1 \left[\beta^{\text{FreeNight}}/30 < 0.119 \right] + 21 \cdot 1 \left[\beta^{\text{FreeNight}}/30 < 1.52 \right] + 1 \left[\beta^{\text{FreeNight}}/30 > 1.52 \right] + 3 \cdot 1 \left[\beta^{\text{FreeNight}}/30 > 0.119 \right] + 9 \cdot 1 \left[\beta^{\text{FreeNight}}/15 > 0.667 \right] + 13 \cdot 1 \left[\beta^{\text{FreeNight}}/15 < 0.667 \right],$$

where all the plans are from T-Mobile, and the comparisons on the first row are between the regional and the two nearest national plans and the second row is the comparison of the two 1500 minute plans. The numbers 15 and 30 are 1500 and 3000 minutes per month divided by 100 cents per dollar, which converts the monthly value into per minute values for each plan. The objective function is maximized at 53 inequalities at $\beta^{\text{FreeNight}}/30 < 0.119$, or $\beta^{\text{FreeNight}} < \3.57 a month. This explains our upper bounds in Table 2.

6.4 Interpreting the WTP for National Coverage

As our conservative lower bound is uninformative, we will focus on our tight bound, which comes from comparing each regional plan to two national plans with similar numbers of minutes from the same carrier. An industry trade group, the CTIA–The Wireless Association, estimates there were more than 1.4 trillion wireless minutes used in the United States in 2005. Multiplying 1.4 trillion minutes by the tight lower bound for the willingness to pay for national coverage of 0.648 cents per minute yields a national willingness to pay of \$9.1 billion. The CTIA reports the 2005 subscriber revenue of US carriers from their 208 million subscribers is \$113.5 billion. \$9.1 billion is therefore 8% of annual subscriber revenue. The \$9.1 billion is a lower bound; our tight upper bound of 0.766 cents per minute comes to \$10.7 billion annually, or 9.4% of the \$113.5 in annual industry revenue.

On an individual basis, the most popular plans in our data are two 450 minute plans that each have a monthly fee of \$40. For this type of plan, at our lower bound the monthly willingness to pay for national coverage is $0.648/100 \cdot 450$, or \$2.92. Our lower bound for the willingness to pay for national coverage is equivalent to 7.2% of the price of a plan.

We interpret our willingness to pay estimates in terms of the actual plan details. Verizon regional plans charge travelers 69 cents a minute. At the lower bound, a consumer will be indifferent to paying a roaming fee of 69 cents a minute and buying a national plan that costs an extra 0.648 cents a minute if 0.9% of the consumer's minutes are for calls made while traveling. This low percentage is consistent with the overall pattern that national plans are much more popular than regional plans.

There are at least two reasons why our estimates of an annual willingness to pay of \$9.1–10.7 billion are overestimates. Section 5.3 examined data from a survey of consumers and found that Amazon users take more trips than non-Ama-zon mobile phone users. Amazon users may value national coverage more. Second, our

data contain customers who purchased only postpaid / monthly contracts. We do not observe customers using prepaid plans, some of which have national coverage, and some of which do not. As many prepaid customers have poor credit or use their phones infrequently, they may value national coverage less than those in our sample.

There are more national plans than regional plans. More consumers may choose an individual regional plan than an otherwise similar national plan simply because an individual national plan has closer substitutes in product characteristic space. This misspecification biases our findings to estimating a low willingness to pay for national coverage and therefore reinforces our finding of a high valuation for national coverage.

7 Conclusions

There has been a tremendous amount of consolidation in the wireless service industry. One motivation is to offer customers seamless national coverage areas, where features such as data and voicemail will work without interruptions. This paper proposed semiparametric demand estimators for market share rank data to estimate the willingness to pay for national coverage. Three plan nests, T-Mobile, Verizon and Verizon family, offer both national and regional plans in the same geographic markets. We formally estimated the willingness to pay for national coverage, and found estimates between 0.648 and 0.766 cents per minute. The lower bound corresponds to 7.2% of the monthly bill of a customer with a popular 450 minute plan.

We extrapolated our bounds from the Amazon sample to the entire US population of 208 million mobile phone customers. We found that the annual consumer value from national calling plans is between \$9.1–10.7 billion. These bounds are 8–9.4% of the the industry’s annual revenue of \$113.5 billion.

We interpret our results as showing that providing national coverage areas is very valuable. To the extent that across-market mergers are necessary to provide national calling areas, the evidence from customer behavior on Amazon supports the view that across-market mergers are efficiency enhancing. Interfirm roaming agreements are another alternative to mergers. Evidence from past industry data and current industry news reports suggests roaming agreements typically involve high per-minute transfers between the home carrier of the traveler and the carrier providing the coverage. Our high estimates of the willingness to pay for national coverage are consistent with the behavior of firms in our industry. Just before our data collection began, Cingular discontinued its regional plans. As we write, large national carriers offer mainly national plans and do not price discriminate against travelers within their native calling areas.

Our econometric contribution is to extend semiparametric demand estimation to the use of market-level data. In particular, we show how to estimate a willingness to pay using only market share ranks. We hope our estimator will encourage others to use online retailers such as Amazon as an easily accessible source of data for demand estimation. We acknowledge that the lack of quantity data requires potentially stronger assumptions for identification. However, data from online retailers have potential advantages. In online data, the economist observes the exact information about the product presented to the consumer by retailers. Online data are also freely available for product categories, such as calling plans, where other high quality data sources may be lacking.

A Proofs

A.1 Proof of Lemma 1

Both ε_{ijm} and ξ_{ijm} are unobservable from the viewpoint of an econometrician. Work with the composite error term $\mu_{ijm} = \varepsilon_{ijm} + \xi_{ijm}$. By Assumption 1, the μ_{ijm} 's are i.i.d. within a nest and independent across nests. In what follows, drop the indices i and m , and use the shorthand notation a_j for $x'_{jm}\beta - p_{jm} + v_{ihm}$.

The total payoff for choice j in nest h is $a_j + \mu_j$. The decision rule in equation (1) becomes

$$\mu_l < a_j - a_l + \mu_j$$

for all choices l . Let $f_h^\mu(\mu_j | \beta, \vec{x}_m, \vec{v}_{im})$ be the density of a μ_j in nest h .

First prove the ‘‘only if’’ direction. By the definition of a choice probability,

$$P(j | \beta, \vec{x}_m, \vec{v}_{im}) = \int_{-\infty}^{\infty} \left(\left\{ \int_{-\infty}^{a_j - a_k + \mu_j} f_h^\mu(\mu_k | \vec{x}_m, \vec{v}_{im}) d\mu_k \right\} \cdot \prod_{l=1, l \neq j, k}^J \left\{ \int_{-\infty}^{a_j - a_l + \mu_j} f_{h(l)}^\mu(\mu_l | \vec{x}_m, \vec{v}_{im}) d\mu_l \right\} \right) \cdot f_h^\mu(\mu_j | \vec{x}_m, \vec{v}_{im}) d\mu_j,$$

where $h(l)$ is a convenience function returning the nest of choice l . If j is in the same nest as k , as in the statement of the lemma, $P(k | \beta, \vec{x}_m, \vec{v}_{im})$ is the same function as $P(j | \beta, \vec{x}_m, \vec{v}_{im})$, except that a_k replaces a_j in the upper limits, and a_j replaces the one term where a_k enters in $P(j | \beta, \vec{x}_m, \vec{v}_{im})$. Let $H(a_j, a_k)$ be $P(j | \beta, \vec{x}_m, \vec{v}_{im})$ as a function of a_j and a_k .

As a_j enters only upper limits of integrals in $H(a_j, a_k)$, $H(a_j, a_k)$ is strictly increasing in a_j . Also, a_k enters negatively in only one upper limit in $H(a_j, a_k)$. Because $f_h^\mu(\mu_k | \vec{x}_m, \vec{v}_{im})$ has full support, by Assumption 1, $H(a_j, a_k)$ is strictly increasing in a_j and strictly decreasing in a_k . Then if $a_j > a_k$, as in the statement of the lemma, $H(a_j, a_k) > H(a_k, a_j)$.

Likewise, the ‘‘if’’ direction is proved as the only way $H(a_j, a_k) > H(a_k, a_j)$ is if $a_j > a_k$.

A.2 Proof of Lemma 2

Drop the m index for simplicity. We are comparing products j and k , which are in the same nest. The rank of product j with a finite sample of I customers is \hat{r}_j . The rank orders the (unobserved) market shares s_j . The condition that $\hat{r}_j > \hat{r}_k$ can be rewritten as $\hat{s}_j > \hat{s}_k$. By dividing though by a positive number, the inequality is also

$$\frac{\hat{s}_j}{\hat{s}_j + \hat{s}_k} > \frac{\hat{s}_k}{\hat{s}_j + \hat{s}_k}.$$

Define \tilde{s} to be $\frac{\hat{s}_j}{\hat{s}_j + \hat{s}_k}$, leaving $1 - \tilde{s}$ to be $\frac{\hat{s}_k}{\hat{s}_j + \hat{s}_k}$. We want to show that $P(\tilde{s} > \frac{1}{2}) > P(\tilde{s} < \frac{1}{2})$.

First consider the case without product market error terms ξ_j . We also condition on I_{jk} , the number of people who buy either j or k . An individual i prefers j over k with probability q_i . By Lemma 1, $q_i > \frac{1}{2}$. However, because the density of errors and the realization of fixed effects vary across consumers, q_i will be different for each consumer. We work with the number (rather than the fraction) of consumers who buy j out of the group who buy either j or k . Call this number r_j . The random variable $r_k = I_{jk} - r_j$ is the number of

consumers who pick k over j . If I_{jk} is even, let $r^* = \frac{I_{jk}}{2}$. If I_{jk} is odd, let $r^* = \frac{I_{jk}}{2}$ as well. We want to show $P(r_j > r^*) > P(r_k > r^*)$.

Now, if $q_i = \frac{1}{2}$ for all i , $P(r_j > r^*) = P(r_k > r^*)$ by the properties of the binomial distribution. By a monotonicity arguments, increasing even one q_i will raise $P(r_j > r^*)$ and consequently weakly lower $P(r_k > r^*)$, as for odd I_{jk} $P(r_j > r^*) + P(r_k > r^*) = 1$ and for even I_{jk} $P(r_j > r^*) + P(r_k > r^*) = 1 - P(r_k = r_j)$. So $P(r_j > r^*) > P(r_k > r^*)$, and the ‘‘only if’’ direction of the lemma is proved. The ‘‘if’’ direction just reverses the above steps, as the only way one product is ranked higher than another more frequently is when the first product has a higher payoff.

The above argument conditioned on a value of I_{jk} . As the lemma holds for any value of I_{jk} , it holds unconditionally as well.

Now consider the case with both ξ_j and ε_{ij} errors. For each realization of ξ_j , each consumer has a q_i that involves the remaining uncertainty over the ε_{ij} terms. Even if j has a higher mean payoff than k , it could be that $q_i < \frac{1}{2}$ because of the realization of ξ_j and ξ_k . The probability q_i is the probability of picking j over k given that the payoff of j is $x'_j\beta - p_j + \xi_j + v_{ih}$ and the payoff of k is $x'_k\beta - p_k + \xi_k + v_{ih}$. By Lemma 1, $q_i > \frac{1}{2}$ when $x'_j\beta - p_j + \xi_j + v_{ih} > x'_k\beta - p_k + \xi_k + v_{ih}$. In other words, either the realization of product and market specific shocks is such that $q_i > \frac{1}{2}$ for everyone or $q_i \leq \frac{1}{2}$ for everyone. So the lemma being proved holds if $x'_j\beta - p_j + \xi_j > x'_k\beta - p_k + \xi_k$ more than half of the time when $x'_j\beta - p_j > x'_k\beta - p_k$, which is does because Assumption 1 states that the ξ 's are i.i.d.

A.3 Proof of Lemma 3

Our identification under sampling error argument will show that the probability limit of $Q_M^2(\beta)$ is uniquely maximized by the parameter vectors in the identified set B^0 . Use the notation r_j for the underlying random variable for market share ranks. If $M \rightarrow \infty$, the maximum score objective function converges to the population objective function

$$Q_\infty^2(\beta) = \sum_{h=1}^H \sum_{j=1}^{J_h} \sum_{k=1, k \neq j}^{J_h} 1 \left[x'_j\beta - p_j > x'_k\beta - p_k \right] \cdot E \{ 1 [r_j > r_k] \}, \quad (9)$$

where we have factored the fixed-across-markets product characteristics out of the expectation over the preferences of customers and the number of such customers in each market m . The probability limit can be rewritten to focus on unique pairs of products as

$$Q_\infty^2(\beta) = \sum_{h=1}^H \sum_{j=1}^{J_h} \sum_{k=j+1}^{J_h} \left\{ 1 \left[x'_j\beta - p_j > x'_k\beta - p_k \right] \cdot E \{ 1 [r_j > r_k] \} + 1 \left[x'_k\beta - p_k > x'_j\beta - p_j \right] \cdot E \{ 1 [r_k > r_j] \} \right\}.$$

For each pair of products, the objective function is the sum of two probabilities times mutually exclusive inequalities. $Q_\infty^2(\beta)$ is maximized if the inequality multiplying the (weakly) greater of $E \{ 1 [r_j > r_k] \}$ and $E \{ 1 [r_k > r_j] \}$ is set to 1. Lemma 2 shows that $E \{ 1 [r_j > r_k] \}$ is larger than $E \{ 1 [r_k > r_j] \}$ precisely when $x'_j\beta - p_j$ is greater than $x'_k\beta - p_k$. Therefore, $Q_\infty^2(\beta)$ is maximized for any $\beta \in B^0$, the identified set.

For identification under sampling error, we also need to show that parameter vectors that are not part of the identified set do not maximize the objective function. Equivalently, we need to prove that if some β maximizes $Q_\infty^2(\beta)$ then $\beta \in B^0$. If β maximizes $Q_\infty^2(\beta)$, the larger of $E \{ 1 [r_j > r_k] \}$ and $E \{ 1 [r_k > r_j] \}$ enters the objective

function for each pair of choices. That term multiplies one of the mutually exclusive indicator functions in β , so this β must maximize the non-sampling error objective function, equation (2). So by the definition of B^0 and Lemma 2, $\beta \in B^0$ and the identified set comprises the maximizers of the probability limit of the objective function under sampling error.

References

- Abrevaya, Jason**, “Rank estimation of a generalized fixed-effects model,” *Journal of Econometrics*, 2000, 95, 1–23.
- Allenby, Greg M. and Peter E. Rossi**, “Marketing models of consumer heterogeneity,” *Journal of Econometrics*, 1999, 89, 57–78.
- Andrews, Donald W. K., Steven Berry, and Panle Jia**, “Confidence Regions for Parameters in Discrete Games with Multiple Equilibria, with an Application to Discount Chain Store Location,” 2005. Yale University Working Paper.
- Bajari, Patrick and C. Lanier Benkard**, “Demand Estimation With Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach,” *The Journal of Political Economy*, 2005, 113 (6), 1239–1276.
- Beresteanu, Arie and Francesca Molinari**, “Asymptotic Properties for a Class of Partially Identified Models,” June 2006. working paper.
- Berry, Steven**, “Estimating Discrete Choice Models of Product Differentiation,” *RAND Journal of Economics*, 1994, 25, 242–62.
- , **James Levinsohn, and Ariel Pakes**, “Automobile Prices in Equilibrium,” *Econometrica*, 1995, 63 (4), 841–90.
- Brynjolfsson, Erik, Yu (Jeffrey) Hu, and Michael D. Smith**, “Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers,” *Management Science*, November 2003, 49 (11), 1580–1596.
- Chernozhukov, Victor and Han Hong**, “A MCMC Approach to Classical Estimation,” *Journal of Econometrics*, 2003, 115 (2), 293–346.
- , —, **and Elie Tamer**, “Inference on Parameter Sets in Econometric Models,” 2005. MIT Working Paper.
- Chevalier, Judy and Austan Goolsbee**, “Price Competition Online: Amazon Versus Barnes and Noble,” *Quantitative Marketing and Economics*, June 2003, 1 (2), 203–222.
- Chintagunta, Pradeep K., Dipak C. Jain, and Naufel J. Vilcassim**, “Investigating Heterogeneity in Brand Preferences in Logit Model for Panel Data,” *Journal of Marketing Research*, November 1991, 28 (4), 417–428.

- Crawford, Gregory S. and Matthew Shum**, “The Welfare Effects of Endogenous Quality Choice: The Case of Cable Television,” February 2006. working paper.
- Fox, Jeremy T.**, “Consolidation in the Wireless Phone Industry,” 2005. NET Institute Working Paper 05-13.
- , “Semiparametric Estimation of Multinomial Discrete Choice Models using a Subset of Choices,” 2006. University of Chicago Working Paper.
- Ghose, Anindya and Arun Sundararajan**, “oftware Versioning and Quality Degradation? An Exploratory Study of the Evidence,” 2006. working paper.
- Imbens, Guido and Charles Manski**, “Confidence Intervals for Partially Identified Parameters,” 2003. University of California–Berkeley Working Paper.
- Kim, J. and D. Pollard**, “Cube Root Asymptotics,” *The Annals of Statistics*, 1990, 18, 191–219.
- Manski, Charles**, “Maximum score estimation of the stochastic utility model of choice,” *Journal of Econometrics*, 1975, 3 (3), 205–228.
- , “Identification of Binary Response Models,” *Journal of the American Statistical Association*, 1988, 83 (403), 729–738.
- Miravete, Eugenio J. and Lars-Hendrik Röller**, “Competitive Nonlinear Pricing in Duopoly Equilibrium: The Early U.S. Cellular Telephone Industry,” April 2004. working paper.
- Nevo, Aviv**, “Measuring Market Power in the Ready-to-Eat Cereal Industry,” *Econometrica*, March 2001, 69 (2), 307–342.
- Pakes, Ariel, Jack Porter, Katherine Ho, and Joy Ishii**, “Moment Inequalities and Their Application,” April 2005. working paper.
- Petrin, Amil**, “Quantifying the Benefits of New Products: The Case of the Minivan,” *The Journal of Political Economy*, 2002, 110 (4), 705–729.
- Rosen, Adam M.**, “Confidence Sets for Partially Identified Parameters that Satisfy a Finite Number of Moment Inequalities,” January 2006. working paper.
- Seim, Katja and V. Brian Viard**, “The Effect of Entry and Market Structure on Cellular Pricing Tactics,” November 2004. working paper.
- Shaikh, Azeem**, “Inference for Partially Identified Econometric Models,” 2005. working paper.