

IDENTIFICATION IN ASCENDING AUCTIONS, WITH AN APPLICATION TO DIGITAL RIGHTS MANAGEMENT

JOACHIM FREYBERGER AND BRADLEY J. LARSEN

ABSTRACT. This study provides new identification and estimation results for ascending (traditional English or online) auctions with unobserved auction-level heterogeneity and an unknown number of bidders. When the seller's reserve price and two order statistics of bids are observed, we derive conditions under which the distributions of buyer valuations, unobserved heterogeneity, and number of participants are point identified. We also derive conditions for point identification in cases where reserve prices are binding and present general conditions for partial identification. We propose a nonparametric maximum likelihood approach for estimation and inference. We apply our approach to the online market for used iPhones and analyze the effects of recent regulatory changes banning consumers from circumventing digital rights management technologies used to lock phones to service providers. We find that buyer valuations for unlocked phones dropped by 39% on average after the unlocking ban took effect, from \$231.30 to \$141.50.

Keywords: Ascending auctions, nonparametric identification, unobserved heterogeneity, unknown number of bidders, sieve maximum likelihood, digital rights, Digital Millennium Copyright Act, grey-market activity, smartphone unlocking

Date: December 23, 2020.

We thank Dominic Coey, Dan Quint, Yoshi Rai, Caio Waisman and researchers at the Federal Communications Commission for helpful comments. We also thank Sharon Shiao, Caio Waisman, and Chuan Yu for research assistance. This project was supported by NSF Grant SES-1530632.

Freyberger: University of Bonn; freyberger@uni-bonn.de.

Larsen: Stanford University and NBER; bjarsen@stanford.edu.

1. INTRODUCTION

This paper presents an approach to jointly solve two identification challenges to empirical auctions work in ascending auctions: unobserved heterogeneity at the auction level and an unknown number of bidders. Unlike sealed-bid auctions, ascending auctions—both traditional English auctions as well as online auctions—proceed sequentially, and some potential bidders planning to place a bid are not observed doing so. Hence the number of bidders (N)—a key element for identification arguments in empirical auctions methods—is often unobserved to the researcher. Previously proposed solutions to this problem of unknown N rely on the assumption of independent, private values (IPV), and consequently little empirical work exists in English or online auctions with unknown N outside of the IPV framework. The IPV framework does not allow for bidder’s values to be correlated through unobserved heterogeneity in the auctioned items, but such unobserved heterogeneity is common in practice.¹ Previous research has suggested solutions to this challenge of unobserved heterogeneity, but these methods require that the researcher observes N and, furthermore, while useful in first price auctions (where often all bids are observed by the researcher) these methods do not immediately apply to English or online auctions given the incompleteness of bid data in ascending auctions (where the researcher rarely observes the thresholds at which the highest-value player would drop out of the bidding). In this paper, we provide a unified framework for nonparametric identification and estimation when both problems exist. In particular, we derive conditions for point identification of the distributions of bidder valuations, unobserved heterogeneity, and the number of bidders, as well as partial identification results when these conditions are not met.

We build on the identification arguments of Song (2004), who suggested an approach to handling settings where the number of bidders is unknown and the researcher observes at least two order statistics of bids in English or online auctions. The Song (2004) approach relies on the assumption that bidders have independent

¹For example, in online auctions, listings often contain pictures and detailed descriptions about characteristics of the items sold that both the seller and potential buyers can observe, but such information is difficult for the econometrician to quantify. Therefore, items in different auctions can differ dramatically in ways that are observable to the seller and bidders but not to the econometrician.

private values, in which case the density of a higher order statistic conditional on a lower order statistic will not depend on N . We demonstrate that the same argument holds even if bidder valuations are only independent *conditional* on auction-level heterogeneity that is unobserved by the econometrician. We also demonstrate that the distribution of the number of bidders is identified.

To nonparametrically identify the distribution of unobserved heterogeneity, we use a similar approach to Li and Vuong (1998), Li et al. (2000), and Krasnokutskaya (2011), relying as they do on the deconvolution result of Kotlarski (1967). These approaches require that the researcher observes two bids that are independent conditional on auction-level unobserved heterogeneity. This approach has been applied to first price auctions in a number of papers (see, for example, Decarolis 2017, Krasnokutskaya and Seim 2011, and others). Various studies (Athey and Haile 2002; Athey et al. 2011; Aradillas-López et al. 2013), however, have highlighted that the deconvolution approach to unobserved heterogeneity cannot be applied to English or online auctions using bids alone because bids represent order statistics and not all order statistics are observed, leading to correlation in the observed bids even when individual valuations represent independent draws from the same underlying distribution.

Our approach circumvents this issue of correlated order statistics by relying on an alternative measure of unobserved heterogeneity available to the researcher in many settings. Specifically, we rely on sellers' reserve prices reported to the auction platform. We demonstrate that when reserve prices are secret or non-binding, the distributions of unobserved heterogeneity and buyer valuations are nonparametrically point identified. When reserve prices are public, this can introduce correlation between reserve prices and observed bids, as bids are only recorded if they exceed the public reserve price. We demonstrate how these binding reserve prices affect the likelihood of observed bids and we derive support conditions under which we still obtain point identification. When these conditions are not met, our results yield partial identification. The data requirements for all of our identification arguments, in which the researcher is concerned with both unobserved heterogeneity and an unknown number of bidders, are the following: 1) the econometrician observes the seller's reserve price;

2) if reserve prices are secret, the econometrician observes at least two order statistics of bids; and 3) if reserve prices are public, the econometrician observes two order statistics of bids if these exceed the reserve price.

We apply these identification arguments, using a maximum likelihood approach, to study the impact of recent legislation regarding consumers circumventing digital rights management. Digital rights management refers to technological locks restricting how consumers use software or hardware. These digital locks are used in computer software, e-books, music, film, cell phones, and in many other products. The US Digital Millennium Copyright Act (DMCA) bans circumvention of these digital locks or production of technologies intended to aid consumers in circumventing digital locks. However, tools and tips for how to circumvent digital locks are readily available on the Internet, and punishment mechanisms for violators of these laws are not necessarily salient to consumers. Therefore, it is unclear whether the DMCA or related legislation has any effect in practice on market primitives, such as consumers' willingness to purchase potentially illegally tampered products. Using data from eBay auctions of used iPhones, we analyze the impact of a recent regulatory change banning smartphone unlocking on bidder valuations for unlocked phones. The application provides insights into this previously unstudied question. In particular, we find that buyer valuations for unlocked smartphones decrease after the ban and the number of bidders increases. The estimated difference in the means of the distributions of buyer valuations for unlocked smartphones in the pre and post periods corresponds to a decrease in the dollar valuation for the phones of about 39% on average, from \$231.30 to \$141.50. This difference suggests that the regulatory change may indeed have had real effects on consumers' willingness to engage in potentially shady behavior.

eBay auctions, such as those that we study in our application, represent an important setting where our methodology can be applied. Other ascending-like auction settings where our data requirements may potentially be satisfied are wholesale used-car auctions (Larsen 2019), Copart virtual auto auctions (Grether et al. 2016), government surplus/sales auction platforms (e.g., GovDeal, PropertyRoom, Municibid,

or IRSAuctions), and other online consumer auction platforms (e.g., TopHatter, eBid, ShopGoodwill, Webstore, AuctionZip, Catawiki, or BidForWine).

RELATED LITERATURE

Earlier work by Athey and Haile (2002) studied identification in ascending auction settings (and not estimation) and demonstrated that, in a symmetric, independent private values setting with no unobserved auction-level heterogeneity, the underlying distribution of bidder valuations is identified from the distribution of an order statistic of valuations and from knowledge of the number of bidders. Using rationality assumptions to partially characterize equilibria, Haile and Tamer (2003) relaxed one of the assumptions of Athey and Haile (2002), allowing for the possibility that order statistics of bids do not necessarily correspond to order statistics of valuations. Haile and Tamer (2003) provided identification as well as estimation arguments that exploit these rationality assumptions to obtain bounds on the distribution of the number of bidders and bounds on objects of interest, such as reserve prices.

This earlier work did not allow for any possibility of unobserved auction-level heterogeneity. Such heterogeneity, from the econometrician's perspective, introduces correlation in bidders' private valuations, and is a special case of the conditionally independent private values model and, more generally, affiliated private values model. Athey and Haile (2002) demonstrated a non-identification argument for the most general affiliated private values model in ascending auctions: valuations are not identified from ascending auction bids (unlike in the case of first price auctions) because the highest-valuation is never observed in the data. In our paper, we do not attempt to solve this more general identification problem (the affiliated private values case), but we do propose a solution to the special case of identification and estimation with unobserved auction-level heterogeneity. To do so, we require that the econometrician observe an additional piece of data not required by Athey and Haile (2002) or Haile and Tamer (2003): the seller's reserve price. With this stronger data requirement we are able to relax the no-unobserved-heterogeneity assumption from earlier work.²

²Our paper illustrates the use of variation in potentially binding reserve prices to obtain partial identification. Athey and Haile (2007) discussed several alternative uses of binding reserve price

We also relax the assumption of Athey and Haile (2002) or Haile and Tamer (2003) that the number of bidders is observed by the econometrician, and again we do so by imposing a stronger data requirement. Specifically, we require that the econometrician observe two order statistics of valuations, not just one. And we require that these bids are equal to the respective players' valuations, embracing the assumption that Haile and Tamer (2003) were able to avoid. Thus our results are not strictly stronger or weaker than previous work, but rather a different set of assumptions that yield identification in a previously unstudied environment (ascending auctions with both unobserved heterogeneity and an unknown number of bidders).

Our arguments for handling an unknown number of bidders build directly on arguments in Song (2004), which Hortacsu and Nielsen (2010) argued has long been “the standard to beat in the empirical online auctions literature” due to its distinct ability to handle an unknown number of bidders. The model does so by relying on the insight of Song (2004) that the distribution of one order statistic conditional on a lower order statistic does not depend on the number of bidders, and we, like Song (2004), provide nonparametric identification and estimation arguments for the distribution of bidders' valuations. Our approach is more general than that of Song (2004), however, in that we allow for unobserved auction-level heterogeneity.³ Other existing work that

information. Recent work by Decarolis (2017) applied the Krasnokutskaya (2011) approach using the reserve price and transaction price from first price auctions as two measurements of unobserved heterogeneity, focusing on a sample in which reserve prices were nearly always non-binding in order to avoid the issue of correlation between bids and reserve prices that we address in this paper. Roberts (2013) exploited reserve prices in a very different fashion, presenting a control function approach for settings in which the reserve price is monotonic in the unobserved heterogeneity; this monotonicity assumption is not appropriate in our setting, where we allow for the possibility that reserve are chosen by sellers with privately known valuations for the good.

³In other work building on the Song (2004) approach, Kim and Lee (2014) developed a test of the independent private values assumption using multiple order statistics. Other approaches exist for handling an unknown number of bidders in first price auction settings, such as An et al. (2010), who demonstrated identification when the econometrician observes an instrument for the number of potential bidders (not all of whom necessarily place bids). Hu et al. (2013) extended these results to apply to settings with non-separable unobserved auction-level heterogeneity (where the number of potential bidders in An et al. (2010) can be considered a form of unobserved heterogeneity in their model) when three bids are observable in first price auctions. Additional work studying unobserved heterogeneity in first price auctions includes Armstrong (2013), Balat (2015, 2016), Seifert and Hüttel (2020), and Andreyanov and Caoui (2020). As explained above, existing deconvolution approaches (Li and Vuong 1998; Li et al. 2000; Krasnokutskaya 2011) have thus far been applied primarily in first price auctions (with a known number of bidders) where, unlike ascending auctions, independent bids are available.

allows for an unknown number of bidders includes Platt (2017) and Hickman et al. (2016), who provided identification and estimation approaches for the distributions of valuations in online auctions with independent private values and no unobserved heterogeneity with an unknown number of bidders, obtaining identification by exploiting the arrival order of bidders.⁴ Adams (2007) proved identification of the valuation distribution in independent private values settings with no unobserved heterogeneity when the distribution of bidders is known. Two recent papers, Mbakop (2017) and Luo and Xiao (2019), obtained identification arguments for ascending auctions *with* unobserved heterogeneity by exploiting cases where the econometrician observes more than two bids. Hernández et al. (2020) used an alternative argument that allows for both unobserved heterogeneity and for some forms of limited information about the number of bidders, such as it being partially observed or the econometrician observing auctions under at least two different known probability distributions for the number of bidders.⁵

Our application contributes to a small literature on digital rights management and copyright infringement (e.g. Stallman 1997; Liu et al. 2003; Walker 2003; Von Lohmann 2004) and the literature on piracy and copyright enforcement more broadly (Harbaugh and Khemka 2010). It remains an open question in this literature how effective regulation is at altering consumers' willingness to engage in circumvention. A related paper demonstrating that firms violating digital copyrights may be unaware that they are infringing is Luo and Mortimer (2016), who study infringement of digital photographs.

The specific application of cellphone unlocking relates to a variety of previous studies that have examined the role digital locks play in raising switching costs of consumers. This work has focused on non-US markets, including, among others, Tallberg

⁴Canals-Cerdá and Percy (2013) provided a *parametric* identification result that incorporates unobserved heterogeneity and an unknown number of bidders.

⁵Several recent papers have demonstrated that certain objects of interest, such as bounds on optimal reserve prices, or buyer and seller surplus, are identified in ascending auction settings with correlated private values under the assumption that the number of bidders is known (Aradillas-López et al. 2013; Coey et al. 2017). Unlike these studies, our approach yields estimates of the underlying valuation distributions, which are useful for studying revenue and welfare under counterfactual auction formats.

et al. (2007) (studying Finland), Maicas et al. (2009) (studying Spain), Nakamura (2010) (studying Japan), and Park and Koo (2016) (studying South Korea). Baker (2007) described several costs consumers face when unlocking a phone, including time and monetary costs and potential invalidation of the handset’s warranty. Farrell and Klemperer (2007) described general theoretical arguments for how lock-in practices, such as handset locking, can create inefficiencies and increase firm profits, in particular in settings with network effects such as telecommunications markets. Finally, in focusing our application on smartphones, we contribute to a nascent literature on this industry more broadly. Sinkinson (2014) and Zhu et al. (2015) examined exclusive contracting deals between Apple and AT&T. Fan and Yang (2020) provided a broad study of the welfare effects of product proliferation and competition in the smartphone industry. Kehoe et al. (2018) analyzed dynamic price competition between Apple and Samsung in tablet and smartphone markets, and Yang (2020) analyzed vertical structure and innovation in these markets.

2. IDENTIFICATION

2.1. Introduction of Model. We analyze static, single-unit ascending auctions where bidders have symmetric private values. For each bidder i , we specify the value to take the following form:

$$V_i = X + U_i.$$

In many settings, as in our empirical application, the researcher may prefer to model valuations in a multiplicative form, $e^{V_i} = e^X e^{U_i}$ and work with logs; all our results hold under multiplicative separability as well, but we state the additively separable version here for ease of exposition. The random variable X is independent of U_i for all i and represents a common component through which bidders’ valuations (and the seller’s reserve price, described below) are correlated. This X is observed by bidders and the seller but is unobserved to the econometrician. Let $U_i \sim F_U$ with density f_U , $X \sim F_X$ with density f_X , and $V_i \sim F_V$ with density f_V .

Let N be a random variable with realizations n representing the number of bidders willing to participate in a given auction. We assume N is independent of X and U . Let U^j refer to the j^{th} -highest U in an auction with at least j bidders but unconditional on the actual realization of the number of bidders N . Thus, U^j is the j^{th} -highest order statistic, $U^{j:N}$, where N is a random variable, and U^j is only defined for $N \geq j$.⁶ Our notation here differs slightly from the conventional notation for order statistics precisely because in our setting N itself is a random variable that is unobserved by the econometrician. We use similar notation to denote order statistics of V .

We allow for the seller's reserve price to be either secret or public. To be precise, in this paper, a reserve price is termed to be *public* if the auction is such that only bids exceeding the reserve price are recorded, and a reserve price is termed to be *secret* if the auction is such that bids need not exceed the reserve price in order to be recorded. We specify the seller's reserve price as

$$R = X + W.$$

We assume that W is independent of (X, U, N) . Let $W \sim F_W$ with density f_W and $R \sim F_R$ with density f_R . We do not directly model the seller's valuation or choice of reserve price (nor assume that these reserve prices are optimal), but rather simply assume reserve prices take the above form, as in Decarolis (2017).⁷ Under standard auction rules, optimality of the reserve price combined with an additively

⁶ U^j is the j^{th} -highest U among N bidders, unconditional on the realization of the random variable N , and is thus a draw from the distribution

$$F_{U^j}(u) \equiv \sum_n \Pr(N = n | N \geq j) F_{U^{n-j+1:n}}(u)$$

where $F_{U^{n-j+1:n}}$ is the distribution of the j^{th} -highest bid conditional on $N = n$, which, given that draws of U are i.i.d., is given by the following (see David and Nagaraja 2003):

$$F_{U^{n-j+1:n}}(u) \equiv \left[\sum_{k=n-j+1}^n \binom{n}{k} F_U(u)^k (1 - F_U(u))^{n-k} \right]$$

⁷Roberts (2013) takes a different approach, assuming away the seller-specific term W and assuming instead that R is an unknown monotonic function of X .

(or multiplicatively) separable valuation for the seller would be sufficient conditions for the reserve price to take the form we assume.⁸

We denote order statistics of bids by B^j . We assume that bidders do not play weakly dominated strategies. For simplicity, we also assume there is no minimum bid increment and that bidders incur no cost of making bids. Under these assumptions, and given the private values environment, equilibrium play during the auction will involve each bidder being willing to bid her valuation unless the current bid exceeds that valuation. In a traditional ascending auction, this behavior entails a bidder dropping out of the bidding when the current bid reaches her valuation. In an online eBay-like auction, it can entail repeated, incremental bidding up to a bidder's valuation or placing a proxy bid equal to the bidder's valuation. In either setting, in spite of bidding up to one's value being an undominated strategy, the valuation of a given bidder may be unobserved in auction data if, when the bidder has the opportunity to bid, the current standing bid has already passed her valuation.⁹ For example, if the first- and second-highest-valuation bidders both bid their valuations before the third-highest-valuation bidder, then it will necessarily be true that $B^3 < V^3$. In this scenario, the *observed* third-highest bid may equal some lower order statistic of valuations (such as V^4 , the fourth-highest bidder's valuation), or may equal some other

⁸If the auction rules are such that the highest bidder wins the good if and only if $B^1 \geq R$, paying R when $B^1 \geq R > B^2$ and paying B^2 otherwise, then the optimal reserve price for a seller of value $X + S$, where S is independent of X and U_i for all i , would satisfy

$$\begin{aligned} R &= X + S + \frac{1 - F_V(R)}{f_V(R)} \\ &= X + S + \frac{1 - F_U(R - X)}{f_U(R - X)} \end{aligned}$$

Letting W be the random variable such that $W = S + \frac{1 - F_U(W)}{f_U(W)}$ will yield the form $R = X + W$, as above, with W independent of X and U . A common alternative rule for ascending auctions (e.g. Larsen 2019) with secret reserve prices is that the highest bidder wins the good if and only if $B^2 > R$ and pays B^2 . In this case a seller with value $X + S$ would optimally choose a reserve price of $X + S$, yielding again the form $R = X + W$ above, with $W = S$. Note that our identification arguments will also hold if R is not a reserve price but some other observable feature of the form $R = X + W$ where X and W satisfy the same assumptions as in the reserve price case.

⁹Our assumptions of private values, zero bidding cost/increments, and no weakly dominated strategies imply that the *only* reason order statistics of bids will fail to equal corresponding order statistics of valuations is because of bidders failing to bid their valuations before the current bid exceeds their valuations.

lower bid placed during the bidding process that does not necessarily correspond to any bidder’s valuation.

We will state our identification arguments for the case where the econometrician observes two order statistics of bids that are *exactly* equal to their corresponding valuation order statistics, $B^j = V^j$ and $B^k = V^k$, where $j < k$. The specific assumption required for $B^\ell = V^\ell$ for some integer ℓ is that, for each $\ell' \leq \ell$, at most one of the bidders with valuation higher than $V^{\ell'}$ bids her valuation before the ℓ' -highest-valuation bidder bids her valuation. This condition roughly means that higher-value bidders tend to bid later in the auction and lower-value bidders tend to bid earlier. Several recent papers provide evidence consistent with this condition in eBay auctions (the focus of our application), including Hendricks and Sorensen (2018), Bodoh-Creed et al. (2018), and Coey et al. (2020), and we document similar evidence in Section 4.2. The assumption is not formally testable in our data, however, but we will describe below an approach for evaluating potential violations of this assumption.

Any two order statistics B^j and B^k will suffice for our arguments, but for the majority of the paper we will focus on the second- and third-highest values: $j = 2$ and $k = 3$. We do so for several reasons. First, under the private values assumption, $B^2 = V^2$; it is only B^3 or lower order statistics that will potentially fall below the corresponding order statistic of valuations.¹⁰ Second, we need at least k bidders to be present in all auctions in order for B^k to be defined, and by assuming smaller values for k and j we have more auction observations to work with. Third, with $j = 2$ and $k = 3$, it is simple to describe the conditions under which $B^3 = V^3$ is more likely to hold. Specifically, B^3 will fall below V^3 if either the first-highest-valuation bidder or the second-highest-valuation bidder (the only two bidders who would ever be willing to bid above V^3) bid above V^3 *before* the third-highest-valuation bidder bids her value. By evaluating auctions in which the first- and/or second-highest-valuation bidders arrive within t units of time of the auction closing, and repeating this analysis for

¹⁰The notion that $B^2 = V^2$ is more likely to hold than such equality for lower bids is explained by Athey and Haile (2002) as follows: “...for many ascending auctions, a plausible alternative hypothesis is that bids $B^{n-2:n}$ and below do not always reflect the full willingness to pay of losing bidders, although $B^{n-1:n}$ does (since only two bidders are active when that bid is placed).”

different windows t , the econometrician can analyze the extent to which violations of $B^3 = V^3$ may be affecting results. Performing this type of sensitivity analysis, therefore, requires an additional piece of data representing some notion of when high-value bidders arrived to the auction.¹¹

Our main identification arguments require observing the following for each auction: The reserve price, R , and two order statistics of bids, denoted B^j and B^k . The econometrician does not observe realizations of X , U , W , or N . We demonstrate identification of the distributions of each of these random variables. We describe three main identification results: First, we obtain point identification when reserve prices are secret. Second, we obtain point identification when reserve prices are public and a support condition is satisfied. Third, we obtain partial identification when reserve prices are public and the support condition is not satisfied. We denote the lower bound of the support of any random variable Y by \underline{Y} and the upper bound by \bar{Y} . Notice that $\underline{V^j} = \underline{V^k} = \underline{V}$ and $\bar{V^j} = \bar{V^k} = \bar{V}$ because V^j and V^k are order statistics from the same distribution as V , and similarly for the supports of U , U^j , and U^k . Let $\phi_Y(t)$ denote the characteristic function of a random variable Y .

We summarize the key assumptions from our discussion thus far in Assumption 1.

Assumption 1. (i) For an auction with n bidders, $R = X + W$ and $V_i = X + U_i$ for $i = 1, \dots, n$, where X , W , U_1, \dots, U_n are mutually independent; (ii) bidders are symmetric and have private valuations; (iii) N is independent of X , W , and U_i for all $i = 1, \dots, N$; (iv) $B^j = V^j$ and $B^k = V^k$; and (v) $\Pr(N \geq k) = 1$.

We adopt part (v) of Assumption 1 as a necessary condition for B^k to be defined. However, our model can still rationalize auctions in which fewer than k bids are observed. In the public reserve price setting, these auctions would be interpreted as cases in which at $N \geq k$ bidders arrive to the auction, but fewer than k bidders have valuations above the public reserve price. These auctions with fewer than k

¹¹In our empirical application, we only observe the timing of the highest bid placed by any given bidder, and therefore we treat this as the *only* bid placed by that bidder. If data contains the timing of all bids placed by each bidder, one could define a bidder's first arrival as the timing of the first bid placed by that bidder. This is the approach adopted by Song (2004), who presented a model in which bidders have multiple, randomly arriving bidding opportunities.

bids explicitly enter into the identification arguments and estimation steps in the public reserve price setting. In the secret reserve price setting, (v) can be relaxed, and auctions with fewer than k bids are rationalized by cases where fewer than k bidders arrived at the auction. These auctions with fewer than k bids can then be safely ignored in identification and estimation, and the distribution of N conditional on $N \geq k$ would still be identified.

We also assume the following:

Assumption 2. $E[|B^j| + |B^k| + |R|] < \infty$ and $E[X] = 0$.

Assumption 3. (i) ϕ_W and ϕ_X have only isolated real zeros. (ii) The real zeros of ϕ_{U^j} and ϕ'_{U^j} are disjoint.

Notice that the means of X , W , U , are not identified without a location normalization. To identify the distributions we therefore impose in Assumption 2 that $E[X] = 0$, but normalizing the mean of W instead yields analogous results. The other moment condition in Assumption 2 is a mild regularity condition and Assumption 3 imposes technical conditions on characteristic functions, which are satisfied by standard distributions (Evdokimov and White 2012). These types of conditions are common in models with multiple measurements; see, for example, Li and Vuong (1998).

2.2. Identification with Secret Reserve Prices. In the case of secret reserve prices, when the econometrician observes B^j , B^k , and R , we obtain the following result:

Theorem 1. *Suppose that Assumptions 1-3 hold. Then F_X , F_W , and F_U are identified from the joint distribution of bids B^j and B^k and secret reserve price R . If the number of points of support of N is finite and F_U is continuous, then the distribution of N is identified as well.*

As highlighted in Section 2.1, in the secret reserve price setting, we can relax part (v) of Assumption 1, allowing for the possibility that N may be less than k , but

still maintaining the assumption that we observe $B^j = V^j$ and $B^k = V^k$ (part (vi)) whenever $N \geq k$. We then can simply do our analysis conditional on $N \geq k$ and still identify F_X , F_W , and F_U because X , W , and U are assumed to be independent of N (part (iii)). The distribution of the number of bidders is then identified conditional on $N \geq k$ (that is $P(N = n \mid N \geq k)$ for all $n \geq k$) under the assumptions of Theorem 1.

The formal proof, which is in the appendix, proceeds in three steps. First we use one observed bid and the reserve price to identify the distributions of X and W , which follows from an extension of Kotlarski's Lemma; see Kotlarski (1967) and Evdokimov and White (2012). While the formal arguments are more involved, it is easy to see that the first two moments of X and W are identified from $E[X] = 0$, $E[W] = E[R]$, $var(X) = cov(W, U^j)$, and $var(W) = var(R) - var(X)$. Second, we show that knowledge of the characteristic function of X implies identification of the joint distribution of U^j and U^k . Also notice that since X , W , and U are independent, identification of the marginal distributions is equivalent to identification of the joint distribution. Finally, arguments related to those in Song (2004) then yield identification of the distribution of valuations and the number of bidders. These arguments can also be used to demonstrate that it is generally possible to identify a parametric distribution of N , even if the support is infinite, for example, if N follows a mixture of Poisson distributions.

2.3. Identification with Public Reserve Prices. In the case of public reserve prices, bids will only be observed if they lie above R . Define $D_1 = \mathbf{1}(R > B^j \geq B^k)$, $D_2 = \mathbf{1}(B^j \geq R > B^k)$, and $D_3 = \mathbf{1}(B^j \geq B^k \geq R)$. We assume that the observed data is a random sample from the distribution of $(R, D_1, D_2, D_3, B^j \cdot (D_2 + D_3), B^k \cdot D_3)$ with $D_1 + D_2 + D_3 = 1$. Notice that we therefore assume that $N \geq k$ and that if B^k is not observed, then $B^k < R$. Point identification is still achieved in this case as long as the support of B is greater than that of R in the strong set order, as we state in the following result.

Theorem 2. *Suppose Assumptions 1, 2, and 3(i) hold and that $\bar{R} \leq \bar{B} < \infty$ and $\underline{R} > -\infty$. Then F_X , F_W , and $P(U \leq u_j \mid U \geq u_k)$ for all $u_j, u_k \geq \underline{W}$ are identified from the joint distribution of $(R, D_1, D_2, D_3, B^j \cdot (D_2 + D_3), B^k \cdot D_3)$. Moreover, if in addition $\underline{B} \geq \underline{R}$, then (i) F_U is identified and (ii) the distribution of N is identified if N has finite support and F_U is continuous.*

The intuition for the identification result is as follows. By the additivity assumption and independence, conditioning on $B^j = \bar{B}$ is equivalent to conditioning on $X = \bar{X}$ and $U^j = \bar{U}$. Since W is independent of X and U , it follows that

$$P(R \leq r \mid B^j = \bar{B}) = P(R \leq r \mid X = \bar{X}, U^j = \bar{U}) = P(W \leq r - \bar{X}).$$

Hence, the distribution of W is identified up to a location shift, which is fixed by the assumption that $E[R] = E[W]$. Similarly, the joint distribution of U^j and U^k is identified by considering $P(B^j \geq b_j, B^k \geq b_k \mid R = \underline{R})$. Finally, using similar arguments as in the proof of Theorem 1 we can then show identification of the distributions of X , U , and N . If $\underline{B} < \underline{R}$, however, then only $P(U \leq u_j \mid U \geq u_k)$ for all $u_j, u_k \geq \underline{W}$ is identified, but we cannot point identify $P(U \leq u)$.¹²

3. ESTIMATION AND INFERENCE

In this section we discuss estimation of the unknown densities f_X , f_W , f_U , and f_{U^k} in both the secret and the public reserve price cases using a nonparametric or semiparametric maximum likelihood approach. Our approach can also accommodate, but does not require, estimating $P(N = n)$, which we describe in Section 4.5.

3.1. Estimation and Inference in the Secret Reserve Case. In the secret reserve price case, the likelihood of the joint distribution of B^j , B^k and R can be obtained by first writing

$$P(B^j \leq b_j, B^k \leq b_k, R \leq r)$$

¹²We state the support conditions for the observed random variables, namely the bids and the reserve price. However, notice that $\bar{B} = \bar{V}$, $\underline{B} = \underline{V}$, $\bar{R} \leq \bar{B}$ is equivalent to $\bar{W} \leq \bar{U}$, and $\underline{B} \geq \underline{R}$ is equivalent to $\underline{W} \geq \underline{U}$.

$$\begin{aligned}
&= \int P(B^j \leq b_j, B^k \leq b_k, R \leq r \mid X = x) f_X(x) dx \\
&= \int P(U^j \leq b_j - x, U^k \leq b_k - x, W \leq r - x \mid X = x) f_X(x) dx \\
&= \int P(U^j \leq b_j - x, U^k \leq b_k - x) P(W \leq r - x) f_X(x) dx.
\end{aligned}$$

The first step uses the law of iterated expectations and the remaining steps the independence assumptions. It follows that

$$\begin{aligned}
f_{B^j, B^k, R}(b_j, b_k, r) &= \int f_{U^j, U^k}(b_j - x, b_k - x) f_W(r - x) f_X(x) dx. \\
&= \int f_{U^j | U^k}(b_j - x \mid b_k - x) f_{U^k}(b_k - x) f_W(r - x) f_X(x) dx.
\end{aligned}$$

Notice that $f_{U^j | U^k}(b_j - x \mid b_k - x)$ is a function of f_U only. For example, when $j = 2$ and $k = 3$,

$$f_{U^j | U^k}(b_j - x, \mid b_k - x) = \frac{2(1 - F_U(b_j - x))f_U(b_j - x)}{(1 - F_U(b_k - x))^2}.$$

Denote the data by $Z_t = (B_t^j, B_t^k, R_t)$, where $t = 1, \dots, T$ denotes an auction. Let $\theta_0 = (f_X, f_W, f_U, f_{U^k}) \in \Theta$, where Θ denotes the parameter space. Define the contribution of an individual auction t to the log-likelihood as

$$l_s(\theta_0, Z_t) = \ln \left(\int f_{U^j | U^k}(B_t^j - x \mid B_t^k - x) f_{U^k}(B_t^k - x) f_W(R_t - x) f_X(x) dx \right),$$

where the s subscript on $l_s(\cdot, \cdot)$ denotes the *secret* reserve price case. Thus, given a random sample of T auctions $\{B_t^j, B_t^k, R_t\}_{t=1}^T$, we can estimate θ_0 by maximizing $\sum_{t=1}^T l_s(\theta_0, Z_t)$. This approach allows for nonparametric estimation, but also for flexible parametric and semi-parametric implementations, such as imposing parametric assumptions on the distribution of unobserved heterogeneity and/or the distributions of valuations. In the empirical application we use a series estimator, which we describe in Section 4.5. More generally, there is a large literature on semi- and non-parametric maximum likelihood estimation, which provides general conditions and results for consistency, rates of convergence, and inference (see, among other, Shen

1997, Murphy and Van Der Vaart 2000, Shen and Shi 2005, Chen 2007, Chen et al. 2011, Akerberg et al. 2012, and reference therein).

3.2. Estimation and Inference in the Public Reserve Case. If the reserve price is public we can still derive the log-likelihood function. To do so, define

$$p_1(r) = \frac{\partial}{\partial r} P(R \leq r, D_1 = 1), \quad p_2(r, b_j) = \frac{\partial}{\partial b_j \partial r} P(B^j \leq b_j, R \leq r, D_2 = 1),$$

and

$$p_3(r, b_j, b_k) = \frac{\partial}{\partial b_j \partial b_k \partial r} P(B^j \leq b_j, B^k \leq b_k, R \leq r, D_3 = 1).$$

We show in Appendix B that

$$p_1(r) = \int_{-\infty}^{\infty} \int_{-\infty}^{r-x} F_{U^j|U^k}(r-x | u_k) f_{U^k}(u_k) du_k f_W(r-x) f_X(x) dx,$$

$$p_2(r, b_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{r-x} f_{U^j|U^k}(b_j-x | u_k) f_{U^k}(u_k) du_k f_W(r-x) f_X(x) dx,$$

and

$$p_3(r, b_j, b_k) = \int_{-\infty}^{\infty} f_{U^j|U^k}(b_j-x | b_k-x) f_{U^k}(b_k-x) f_W(r-x) f_X(x) dx.$$

Notice that since $f_{U^j|U^k}$ and $F_{U^j|U^k}$ depend on f_U only, all three expressions only depend on the four densities f_X , f_W , f_U , and f_{U^k} . The data of auction t consists of $\tilde{Z}_t = (R_t, D_{1t}, D_{2t}, D_{3t}, B_t^j \cdot (D_{2t} + D_{3t}), B_t^k \cdot D_{3t})$ and the contribution of auction t to the log-likelihood is $l_p(\theta_0, \tilde{Z}_t) = D_{1t} \ln(p_1(R_t)) + D_{2t} \ln(p_2(R_t, B_t^j)) + D_{3t} \ln(p_3(R_t, B_t^j, B_t^k))$, where p in $l_p(\cdot, \cdot)$ denotes the *public* reserve price case. Again, using a random sample of auctions, we can estimate θ_0 parametrically, nonparametrically or semiparametrically. Notice that our identification arguments in this setting rely on support conditions and on conditioning on random variables on the boundary of the support. It is therefore possible that the estimator, even in a point identified parametric setting, converges slowly (see Khan and Tamer (2010) for a discussion on irregular estimation and support conditions). Moreover, as shown in Theorem 2 and the related discussion, if the support conditions do not hold, then θ_0 may not even be point identified.

In these cases, we can use estimation and inference methods that are robust to potential partial identification (see for example Liu and Shao (2003) and Chen et al. (2018) for flexible parametric results and Chen et al. (2011) for semi- and nonparametric results). These methods yield confidence sets for functionals of θ_0 , such as the means and the variances of the distributions, but they are more complicated to implement. In our empirical application we focus on a setting with secret reserve prices.

4. APPLICATION TO USED SMARTPHONE AUCTIONS

4.1. Background on Digital Rights Management and Smartphone Unlocking. Digital rights management (DRM) refers to technological locks placed on software or hardware to restrict its use or modification. The use of these locks has been highly controversial. Proponents of DRM argue that these restrictions are necessary to prevent *grey-market* activity, such as copyright infringement of digital intellectual property (Liu et al. 2003). Opponents argue that DRM takes a step beyond traditional copyright law by controlling how consumers access or use goods or digital content they have legally purchased, suggesting that these laws instead serve primarily to restrict competition between producers (Von Lohmann 2004; Walker 2003; Stallman 1997). A number of products are controlled through DRM, including, among many others, computer software, with digital locks enforcing limited installs or requiring activation keys; e-books, music, or film, with limits on sharing or on device compatibility; and cellular handsets, with digital locks between the subscriber identification module (SIM) and the phone's software, restricting the handset to only function on a particular provider's cellular service network.

In the United States, the key law regarding DRM is the Digital Millennium Copyright Act (DMCA) of 1998. This law was implemented in response to the 1996 copyright treaty of the World Intellectual Property Organization, which required its members (including the United States) to adopt measures to prohibit tampering with digital locks. In the early years of the DMCA, cellular handset unlocking was granted an explicit exemption, and consumers could legally unlock their out-of-contract phones through a variety of do-it-yourself or third-party services (Van Camp

2013). In late 2012, the copyright office of the Library of Congress failed to renew this exemption, arguing that phone unlocking tampers with copyrighted firmware and hence is arguably in violation of the law (Federal Register 2012; Coutts 2012). This change made phone unlocking illegal as of January 26, 2013, imposing fines ranging from \$200–2,500 for unlocking a phone and, for selling an unlocked phone, fines ranging from the seller’s profits to potentially as high as \$500,000 (and a sentence of five years in prison; Coutts 2013). Contemporary conversations among consumers online suggest that consumers were nervous as to how this massive fine and prison sentence would be enforced and to whom it would apply (Velazco 2013; Khanna 2013). In response to backlash from consumer advocates (Wyatt 2013), a bill was eventually signed into law in 2014 to re-allow consumer unlocking of phones.

Although laws such as the DMCA have arisen to prohibit the production or distribution of technology intended to circumvent these digital locks, these laws and copyright laws in general are notoriously difficult to enforce (Harbaugh and Khemka 2010), violations are difficult to police, and the severity of the punishment is ambiguous to market participants. Given these enforcement challenges, it remains an open question whether these laws are effective in altering individuals’ (buyers and sellers) willingness to engage in grey-market activities like DRM circumvention. We contribute to this question by examining buyers’ willingness to pay and sellers’ pricing for DRM-tampered goods—unlocked smartphones—before and after the January 2013 ban on unlocking. The analysis provides several insights that we discuss below.

4.2. Data and Reduced-Form Analysis. We use a new dataset of eBay auctions for used iPhones.¹³ The sample consists of used iPhone 4 and 4S models with 8, 16, or 32 gigabytes (GB) of memory and of black or white color. These models were the most frequently auctioned iPhone models during our sample period, September 22, 2012 to May 21, 2013. We choose this sample period as it begins after the introduction of the iPhone 5 (September 21, 2012) and includes the date of the regulatory change banning phone unlocking, January 26, 2013. The models we study were released between June 24, 2010 and October 14, 2011, prior to the start date of our sample,

¹³We obtained this data from internal eBay servers through a contract with eBay.

and thus a large number of used (unlocked and locked) iPhone 4 and 4S handsets had accumulated and were being sold in our sample period.

We focus our application on auctions in which the seller uses a secret reserve price. For each auction, the data contains the second and third order statistics of bids, the seller's secret reserve price, and several characteristics of the listing, including an indicator for whether the phone is locked to a particular carrier (Verizon or AT&T) or unlocked. We drop all auctions in which the bids, the reserve price, or the shipping fee lie outside of their respective 0.01 and 0.99 quantiles. We also drop auctions in which the bids or the reserve price lie above the contemporary price of a brand new iPhone 5 (\$649) sold at the Apple Store (see Wyatt 2013). This leaves a sample of 12,872 auctioned iPhones. Descriptive statistics for this sample are found in Table 1. Table 1 demonstrates that 55% of phones in the sample are locked to AT&T, 34% are locked to Verizon, and 11% are unlocked. We find that lower bids tend to arrive earlier in the auction: the average third-highest bid arrives nearly twice as far from the end of the auction as does the average second-highest bid.

In Table 2, we report means and standard deviations (in parentheses) for several variables broken down by unlocked vs. locked phones before and after the ban. We also show the difference between these means and the difference-in-differences estimates (along with standard errors in parentheses). From the final column, we see that the difference-in-differences for reserve prices and the number of bids is not statistically significant. The difference-in-differences for the second-highest bid and the number of bidders are both significant. The results imply that the average transaction price decreases by \$14.9 *less* for unlocked phones than for locked phones. And the number of observed bidders increases by 0.7 more for unlocked phones than for locked phones. These statistics highlight an important reason why a structural model can shed insight in the question of what happened to consumer valuations before and after the ban: transaction prices decrease less for unlocked phones than for locked phones, and this could have been driven either by changes in the *intensive* margin (the underlying willingness to pay of consumers) or the *extensive* margin (the

underlying distribution of N). Both of these objects are *unobservable* in the data, but our approach allows us to tease them apart.¹⁴

4.3. Evaluating Model Assumptions in eBay Used Smartphone Auctions.

As with many real-world ascending-like auctions, the number of bidders N is not observed in our setting: potential bidders who arrive at the auction after the standing bid has passed their values will not be observed placing a bid. These auctions also exemplify a setting in which it is important to account for unobserved heterogeneity, as used smartphones can differ dramatically in ways that are observable to the seller and bidders but not to the econometrician, such as through a cracked screen, scratched surface, missing USB adapter, or faulty battery, or in positive ways, such as a lack of wear and tear.¹⁵ Pictures and detailed descriptions posted by the seller contain information about these characteristics that both the seller and potential buyers can observe, but such information is difficult for the econometrician to quantify.¹⁶

Our model requires the assumption that N is independent of the unobserved heterogeneity. This assumption cannot be tested directly, but we can assess how the number of *observed* bidders correlates with the observed features of the listing. Columns 1–3 of Table 3 present the results of a regression of the observed number of bidders on the log of one plus the number of photos, a set of phone-type fixed effects, or both. We find that these observable features of the listing explain little of the variation in the number of observed bidders (the R^2 in each case is less than .05), and the number of photos is insignificant in each regression. The model also requires the assumption

¹⁴The final two rows of Table 2 demonstrate that the total number of distinct sellers and the total number of listings decreased for both locked and unlocked phones. The decrease in the absolute number of listings was much greater for locked phones, but in relative terms these supply side changes were quite similar in both markets (both markets saw the number of sellers and listings decrease by about 50%).

¹⁵Our focus on used phones is also due to the fact that new unlocked phones are more likely to be legally unlocked by the original vendor, and thus the alteration of the DMCA is less likely to impact new phones.

¹⁶When observable to the econometrician, text descriptions could be exploited using natural language parsing algorithms or images could be analyzed with image processing algorithms, and this could aid in accounting for item-level heterogeneity. In such cases our approach would remain useful to account for remaining unobserved heterogeneity.

that the second- and first-highest-value bidders did not both arrive before the third-highest. We evaluate potential violations of this assumption in Section 4.6 through the sensitivity analysis described in Section 2.1, where we limit the data based on the arrival times of the first- and second-highest bids.

Our identification arguments also require the assumption that bidders are short lived, exiting the auction market after one attempt. In practice, this appears to be a reasonable approximation, as 71% of bidders in our data bid in at most one auction for a given phone specification.¹⁷ We also assume bidders' valuations are private. As with many auction settings, allowing for interdependencies in valuations would be preferable but would be beyond the state of the methodological literature. However, we believe that the private values assumption is a reasonable approximation to reality here in that all buyers have access to the same information on the website about the product.¹⁸ The private values assumption might also be violated if bidders are purchasing phones primarily to resell them, as correlation among bidders' signals of future market conditions could introduce interdependencies in valuations. We find that very few buyers purchase the same product twice.¹⁹

Finally, our model also assumes bidder valuations are independent. This assumption is motivated in our environment by idiosyncratic differences in willingness to pay across bidders that can arise due to differences in ability to pay, outside option opportunities, or urgency with which the phone is needed. For example, Coey et al.

¹⁷We computed this statistic using internal eBay data containing bidder identities; these identities are not contained in our final dataset approved by eBay for estimation. While a non-trivial fraction of bidders *bid* in multiple auctions, very few actually *purchase* multiple items, as described below.

¹⁸Lemons-like interdependencies—arising from the seller withholding information from the buyer about the quality of the good—are less likely to be a concern in our data than in many other auction settings due to buyer protection plans and sanctions against deceptive sellers, which eBay has incorporated in recent years.

¹⁹We cannot examine this directly in our data, which lacks bidder identities, but using similar data from Coey et al. (2020), who studied consumer search using eBay auctions from 2013–2014, we find that only 3.4% of buyer-product pairs involve the same buyer purchasing the same product more than once, and only 0.24% of seller-product pairs involve a seller who also at some point in the sample period was a *buyer* of that same product. Also, using data on eBay fixed-price listing of new iPhones from Kehoe et al. (2018), we find that, over the three 8.5 month time periods (a similar length to our sample) from October 2014 through February 2017, an average of 9.3% of buyers purchase more than one phone.

(2020) demonstrated that eBay auction bidding behavior is consistent with bidders having idiosyncratic, independent deadlines by which they need a particular item.

4.4. Controlling for Observables. Before implementing our structural estimation, we first control for observable auction-level heterogeneity using the standard homogenization step of Haile et al. (2003). We account for shipping fees by adding them to the observed bids and reserve prices. We adopt the log (i.e. multiplicatively separable) specification described at the beginning of Section 2. We perform the homogenization step with the following regression

$$R = Z'\beta + X + W,$$

where R here is the log reserve prices levels, $X + W$ is the residual (X is the log of unobserved heterogeneity level and W is the log of the private component), and Z is a vector of observable characteristics, consisting of the log of one plus the number of photos and a fully saturated model of indicators for all phone types. The resulting $\hat{\beta}$ estimates from this regression are shown in column 6 of Table 3. Columns 4 and 5 display results from controlling for the number of photos or the phone type indicators separately. Unlike columns 1–3, where we find that observable characteristics do not predict well the number of observed bidders, we find that these characteristics do explain a non-trivial portion of the variance in reserve prices (the R^2 is 0.37 columns 5 and 6).²⁰ After evaluating this regression, we compute homogenized bids and reserve prices by subtracting the regression’s predicted value from log bids and log reserve prices. These residuals are then the inputs into our structural estimation exercise.

4.5. Details of Structural Estimation and Inference. We present results for a semiparametric and a nonparametric specification of the model. For the semiparametric model, we specify the marginal densities f_U , f_W , and f_X , to all be normal and we only estimate the means and the variances, except for the mean of X , which we

²⁰The fact that sellers’ reserve prices have a statistically significant relationship with these observable characteristics and explain a non-trivial portion of the reserve price variance is suggestive that reserve prices may also be related to *unobserved* auction-level characteristics, as our model specifies, given that, from the seller’s perspective, there is no difference between observable and unobservable auction-level heterogeneity.

set to 0. We then also nonparametrically estimate f_{U^k} using a series estimator. This flexibility in the k^{th} order statistic density f_{U^k} permits the model to accommodate the unknown distribution of the number of bidders. Specifically, we approximate f_{U^k} using a Hermite polynomial, similar to Gallant and Nychka (1987).²¹ This means that we approximate

$$f_{U^k}(u) \approx \frac{1}{\sigma_{U^k}} \left(\sum_{m=1}^M \theta_m^{U^k} \left(\frac{u - \mu_{U^k}}{\sigma_{U^k}} \right)^{m-1} \right)^2 e^{-\left(\frac{u - \mu_{U^k}}{\sigma_{U^k}} \right)^2},$$

where M is a smoothing parameter; we choose $M = 4$. The parameters to be estimated are then the vector θ^{U^k} , as well as μ_{U^k} , σ_{U^k} , and the means and variances of the normal distributions. The location and scale parameters μ_{U^k} and σ_{U^k} are not required for consistent estimation but improve the performance of the estimator. We estimate them in a parametric initial step, assuming that U^k is normally distributed as well. We then plug in this approximating polynomial with the estimated values into the likelihood expression and maximize over θ^{U^k} and the means and variances of the normal distribution. We perform the integration in the likelihood by Gauss-Hermite quadrature (see Judd 1998).

We also impose the constraint that the density of U^k integrates to 1. This restriction amounts to an equality constraint that is quadratic in $\theta_m^{U^k}$ because

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{\sigma_{U^k}} \left(\sum_{m=1}^M \theta_m^{U^k} \left(\frac{u - \mu_{U^k}}{\sigma_{U^k}} \right)^{m-1} \right)^2 e^{-\left(\frac{u - \mu_{U^k}}{\sigma_{U^k}} \right)^2} du &= \int_{-\infty}^{\infty} \left(\sum_{m=1}^M \theta_m^{U^k} z^{m-1} \right)^2 e^{-z^2} dz \\ &= \left(\theta_m^{U^k} \right)' \Omega_M \theta_m^{U^k} \end{aligned}$$

where Ω_M is an $M \times M$ matrix with element (m, m') equal to

$$\int_{-\infty}^{\infty} z^{m+m'-2} e^{-z^2} dz = \sqrt{\pi} E[Z^{m+m'-2}]$$

and $Z \sim N(0, 0.5)$.

²¹Our estimator differs from most previous studies relying on the result of Kotlarski (1967), which directly estimate a joint characteristic function and then apply inverse Fourier transforms to recover underlying distributions. The maximum likelihood estimation approach allows us to estimate all parameters in one step.

For the nonparametric approach we proceed analogously but also approximate the densities of f_U , f_W , and f_X using Hermite polynomials. For this exercise, we specify f_U , f_W , and f_{U^k} each as fourth-degree Hermite polynomials. For the density f_X we use a third-degree Hermite polynomial.²² We set the scale and location parameters to the estimates from the semiparametric model and maximize over the coefficients of the polynomials. The nonparametric model then has by construction a better fit and is more flexible than the semiparametric model. We can therefore also view the nonparametric estimator as a much more flexible parametric approach. We have to ensure that all densities integrate to 1 (as described above) and now also that the mean of X is 0, which is another quadratic constraint.²³

In addition to point estimates for these distributions, we compute standard errors for their estimated means and standard deviations. To do so, we treat our estimators as flexible parametric estimators. The estimated parameter vector is then asymptotically normally distributed and we can obtain standard errors for features of the distribution using the delta method.

Our estimation strategy does not require estimating $\Pr(N = n)$ in order to estimate f_U , f_{U^k} , f_W , and f_X , but once we have these objects, we can use the relation

$$F_{U^k}(u) = \sum_{n=\underline{n}}^{\bar{n}} \Pr(N = n) \left[\sum_{i=n-k+1}^n \binom{n}{i} F_U(u)^i (1 - F_U(u))^{n-i} \right]$$

to estimate the distribution of the number of bidders by minimizing the L^2 -norm of the difference between the left- and right-hand-sides of this expression, subject to all probabilities being non-negative and summing to 1. For our empirical application, we adopt two specifications for $\Pr(N = n)$. The first is a parametric approximation

²²This latter choice is driven by the fact that X , the unobserved heterogeneity, is the dimension over which we integrate in computing the likelihood, and, even with many Gauss-Hermite nodes, there can be few nodes close to zero. With an even-degree polynomial, this can yield an estimated density with little mass exactly at zero and a mode on either side of zero. With a third-degree polynomial approximation, the integration and estimation result in an estimated density that is single-peaked. The qualitative results and the pre and post comparison are similar when all densities are approximated with fourth-degree polynomials.

²³Using analogous algebra as before, we can write the constraint as $(\theta_m^X)' \tilde{\Omega}_M \theta_m^X = 0$ where $\tilde{\Omega}_M$ is an $M \times M$ matrix with element (m, m') equal to $\sqrt{\pi} E[Z^{m+m'-1}]$ and $Z \sim N(0, 0.5)$.

using a mixture of eight Poisson distributions, given by

$$\Pr(N = n) \approx \sum_{m=1}^Q \alpha_m \left(\frac{\lambda_m^n e^{-\lambda_m}}{n!} \right)$$

where Q is a smoothing parameter and $\{\alpha_m\}_{m=1}^Q$ and $\{\lambda_m\}_{m=1}^Q$ are parameters to be estimated. Here we choose $Q = 8$. The second specification is nonparametric, given by $\Pr(N = n) = \alpha_n$, where $\{\alpha_n\}_{n=3}^{\bar{n}}$ are parameters to be estimated and where we set \bar{n} to be large—ranging from 200 to 10,000 depending on the estimation sample.

4.6. Estimation Results. Using homogenized bids and reserve prices, we estimate the distributions of unobserved heterogeneity, bidder valuations, and reserve prices, where the latter two distributions are net of the unobserved heterogeneity. By Theorem 1, each of these objects (F_U , F_W , and F_X for each carrier-period combination) is point identified. Table 4 displays the model estimates. All units are in log points. Panels A and B display results from the semiparametric model and panels C and D from the nonparametric model (described in Section 4.5). Panels A and C use the full sample and panels B and D limit the sample to auctions in which the first- and second-highest bids arrived within the last five minutes of the auction, which we refer to as the *late-bidding* sample.²⁴

We estimate the model separately for each carrier code (AT&T, Verizon, and unlocked) and separately for each time period (pre vs. post the unlocking ban). For each case, we report the mean and standard deviation of U and W and the standard deviation of X , with standard errors in parentheses. Table 4 also displays the difference between the pre and post estimates, and the difference-in-differences between unlocked and AT&T phones (the second-to-last column) and unlocked and Verizon phones (the final column). All differences or difference-in-differences estimates with t-statistics of magnitude greater than 1.96 are bolded for readability.

²⁴As highlighted in Section 2.1, if both the first and second-highest-valuation bidders bid their valuations before the third-highest-valuation bidder arrives at the auction, it will necessarily be the case that $B^3 < V^3$. When either (or both) of these two-highest-value bidders do not place their highest bid until late in the auction, the chances are higher that $B^3 = V^3$, as required by our identification arguments.

Our semiparametric estimates in panel A suggest that the unlocking ban did not lead to differential changes in F_U , F_W , and F_X for AT&T vs. unlocked phones. For Verizon phones, the estimates suggest that the variance of U increased by 0.078 log points more for unlocked phones than for Verizon phones, and the mean of W increased by 0.058 log points more.²⁵ However, in panel B, when we limit to the late-bidding sample, we find no significant effect of the unlocking ban.

To examine whether any of the above implications are driven by parametric restrictions, we now turn to the results of the nonparametric model. The estimates in panel C show that the mean of buyer valuations $E[U]$ dropped by more for unlocked phones than for AT&T (by 0.435 log points) or for Verizon phones (by 0.554 log points). For unlocked phones the decrease corresponds to a drop of average buyers' willingness to pay for iPhones of 39%, from \$231.30 to \$141.50.²⁶ This significant decrease is robust to estimating the nonparametric model on the late-bidding sample (panel D), where we find similar estimates for the difference-in-differences for $E[U]$. We observe several other significant effects in panel C, but many of these results are not robust to limiting to the late-bidding sample. The decrease in buyer valuations for unlocked phones is robust to this sample restriction—whether compared to AT&T or Verizon phones. We see this decrease as the key finding arising from this nonparametric analysis. This shift is suggestive that buyers may have been less willing to pay for technologically circumvented handsets after the unlocking ban took effect.²⁷

²⁵Throughout Table 4, we find that consumers value Verizon phones less than AT&T phones. This is likely simply an artifact of Apple phone contracts for Verizon being a relatively new phenomenon during our sample period when compared to such contracts for AT&T. We do not report this difference-in-differences comparison to conserve space, and because our primary focus is the unlocking ban.

²⁶These average dollar amounts can be obtained by evaluating valuations at the average predicted value from the homogenization regression, the average unobserved heterogeneity component, and the average e^U : $E[e^{Z'\hat{\beta}}] * E[e^X] * E[e^U]$.

²⁷We note one word of caution in interpreting these results: our estimation exercise constrains the mean of X to be zero in every estimation sample. If, in practice, the mean of X changes before and after the ban, our estimation exercise will misinterpret this as a change in buyer valuations and as a change in the private component of reserve prices, because any non-zero component to $E[X]$ will be absorbed into the mean of W and U . Our estimation strategy does allow other features of the distribution of X to change as long as they do not alter the mean of X . In panel D, we detect a difference before and after the ban in the standard deviation of X , suggesting that, in addition to changes in buyer valuations, unlocked phones may have been different after the ban in terms of

We display the nonparametric results graphically in Figures 1 – 3 for the main sample. Figure 1 includes all of the nonparametric estimates for AT&T, Verizon, and unlocked phones in the pre period, and Figure 2 contains the analogous estimates in the post period. Figure 3 displays a side-by-side comparison of the pre and the post period density estimates for each carrier. These figures suggest that the distribution of unobserved heterogeneity is quite similar for AT&T, Verizon, and unlocked phones, both before and after the unlocking ban. The point estimates for the distribution of seller reserve prices (W) have a clear stochastic ordering in both the pre and post periods, with unlocked phones being highest, then AT&T phones, and then Verizon phones. The ordering for buyer valuations in the pre period is roughly the same, although buyer valuations for AT&T and unlocked phones are quite similar to one another. In the post period, however, we see that the ordering for buyer valuations changes, with unlocked phones being valued much less than AT&T phones, although still more than Verizon phones. Figure 3 demonstrates that the largest visible shift in U , W , X between the pre and post periods occurs for buyer valuations (U) for unlocked phones.²⁸

One important feature of using our approach for this analysis is that it allows us to separately isolate changes in the intensive margin—the *willingness to pay* of bidders (F_U)—from changes in the extensive margin—the *number of bidders* willing to participate (the distribution of N , $\Pr(N = n)$). Changes in either of these objects could lead to changes in overall demand for unlocked vs. locked phones. In Figure 4, we display parametric and nonparametric estimates (using the estimators described in Section 4.5) of the CDF of the number of bidders before and after the unlocking ban. For auctions of AT&T and Verizon phones (the top and middle rows of Figure 4), we find similar estimates of the distribution of the number of bidders before and after the ban: the median N for AT&T auctions was 6 in the pre period and 7 in the post period; for Verizon phones the corresponding medians were 12 and 9. For

the unobserved heterogeneity component; that is, the kinds of unlocked phones being sold may have differed before and after the ban. This difference is insignificant in panels A–C, however.

²⁸Figures for the late-bidding sample are similar, and can be found in the working paper version of this study, Freyberger and Larsen (2019).

unlocked phones, however, we find an increase in the number of bidders (the median N increased from 7 to 46).²⁹

Our results indicate that the total number of bidders seeking unlocked phones *increased* after the unlocking ban, while the underlying valuations *decreased*, suggesting that the additional bidders on unlocked phones after the ban were generally lower-valuation bidders than those bidding prior to the ban.³⁰ These results suggest that regulations intended to reduce grey-market activity may alter not only consumers' willingness to engage in such activity but also alter *where* this grey-market activity takes place. For example, sales platforms directly dedicated to phone sales may have been able to react to the smartphone-unlocking ban with policies to restrict unlocking sales, while eBay did not explicitly do so (as a large digital platform receiving only a small fraction of its revenue from phone sales), which may have led to an increase in unlocked-phone purchasers shifting onto the eBay platform during the unlocking ban.³¹

²⁹These medians correspond to the nonparametric fit for the CDF of N ; the Poisson mixture model yields similar results. We find similar qualitative findings in the late-bidding sample, although there we find more weight on large values of N for all carrier codes. See Freyberger and Larsen (2019) for these results. It is to be expected that the number of observed bidders (which has a mean of 12 and max of 30 in Table 1) would be lower than the true N given that some bidders arrive at the auction but never place a bid. Even so, our results suggest that N can be rather large (with positive probability of $N > 100$ or even 1,000 in some panels. These large estimates arise in part because, for larger N , distributions of higher order statistics (of N *iid* draws of a random variable) do not change much with additional increases in N , so even minor changes in order statistics lead to large implied changes in the distribution of N . This may indicate that there are other dimensions through which valuations are correlated other than the scalar unobserved heterogeneity that we model; accounting for such correlation is beyond the current state of the literature, but might lead to less drastic estimates of the tails of N . If, however, N is viewed as being the number of unique users who *viewed* a given listing, regardless of whether they bid, our estimates of N agree with evidence from other eBay listings. We do not see the number of viewers in our data, but in the publicly available eBay Best Offer listing data of Backus et al. (2020), we find that for popular used cell phone products (those listed at least 500 times) posted from June 2012 to May 2013, the median number of viewers is 32, the 90th percentile is 159, the 99th percentile is 587, and extreme observations have viewers in the 1,000s.

³⁰Note that these insights would have been missed without the structural model: Table 2 shows that auction prices (second-highest bids) and the observed number of bidders both increased for unlocked phones relative to locked phones.

³¹Consistent with this idea that the ban may have shifted where grey-market activity takes place, contemporary online news (Van Camp 2013) reported that US traffic on a phone unlocking website in the UK (where unlocking was never illegal) increased during the period of the US ban.

5. CONCLUSION

This paper introduces a new approach to identification in auctions with unobserved auction-level heterogeneity and an unknown number of bidders. The methodology relies on deconvolution ideas for handling unobserved heterogeneity—which have been applied extensively to first price auctions but not to ascending auctions due the complicating factor of correlation between order statistics. The approach also relies on order statistics comparisons that have previously been limited to settings that do not allow for unobserved heterogeneity. We bring these ideas together in a unified framework, exploiting information contained in reserve prices—either secret or public—chosen by the seller. We provide point identification as well as partial identification results for these settings and propose a nonparametric sieve maximum likelihood approach or semiparametric approach for estimation and inference.

We apply this framework to analyze changes in bidders' willingness to pay before and after 2013 regulatory changes banning the removal of digital locks on cellphones. We find that buyer valuations for unlocked phones decreased and the number of bidders for these phones increased after the ban relative to corresponding changes for phones locked to AT&T or Verizon. These results are suggestive that digital rights management laws, while difficult to enforce in practice, may have real effects on consumers' willingness to pay and purchase activity.

REFERENCES

- Ackerberg, D., Chen, X., and Hahn, J. (2012). A practical asymptotic variance estimator for two-step semiparametric estimators. *Review of Economics and Statistics*, 94(2):481–498.
- Adams, C. P. (2007). Estimating demand from ebay prices. *International Journal of Industrial Organization*, 25(6):1213–1232.
- An, Y., Hu, Y., and Shum, M. (2010). Estimating first-price auctions with an unknown number of bidders: A misclassification approach. *Journal of Econometrics*, 157(2):328–341.
- Andreyanov, P. and Caoui, E. H. (2020). Secret reserve prices by uninformed sellers. Working paper, University of Toronto.
- Aradillas-López, A., Gandhi, A., and Quint, D. (2013). Identification and inference in ascending auctions with correlated private values. *Econometrica*, 81(2):489–534.
- Armstrong, T. B. (2013). Bounds in auctions with unobserved heterogeneity. *Quantitative Economics*, 4(3):377–415.
- Athey, S. and Haile, P. A. (2002). Identification of standard auction models. *Econometrica*, 70(6):2107–2140.
- Athey, S. and Haile, P. A. (2007). Nonparametric approaches to auctions. *Handbook of Econometrics*, 6:3847–3965.
- Athey, S., Levin, J., and Seira, E. (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *Quarterly Journal of Economics*, 126(1):207–257.
- Backus, M., Blake, T., Larsen, B., and Tadelis, S. (2020). Sequential bargaining in the field: Evidence from millions of online bargaining interactions. *The Quarterly Journal of Economics*, 135(3):1319–1361.
- Baker, C. A. (2007). Consumer switching costs in the us marketplace for wireless telephone service. Working paper, AARP Public Policy Institute.
- Balat, J. (2015). Highway procurement and the stimulus package: Identification and estimation of dynamic auctions with unobserved heterogeneity. Working paper, Johns Hopkins University.

- Balat, J. (2016). Identification and estimation of affiliated private values auctions with unobserved heterogeneity. Working paper, Johns Hopkins University.
- Bodoh-Creed, A., Boehnke, J., and Hickman, B. (2018). How efficient are decentralized auction platforms? *Review of Economic Studies*, forthcoming.
- Canals-Cerdá, J. J. and Percy, J. (2013). Arriving in time: Estimation of english auctions with a stochastic number of bidders. *Journal of Business & Economic Statistics*, 31(2):125–135.
- Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In Heckman, J. and Leamer, E., editors, *Handbook of Econometrics*, volume 6 of *Handbook of Econometrics*, chapter 76, pages 5550–5623. Elsevier.
- Chen, X., Christensen, T. M., and Tamer, E. (2018). Monte carlo confidence sets for identified sets. *Econometrica*, 86(6):1965–2018.
- Chen, X., Tamer, E. T., and Torgovitsky, A. (2011). Sensitivity analysis in semiparametric likelihood models. Cowles Foundation Discussion Paper 1836.
- Coey, D., Larsen, B., and Platt, B. C. (2020). Discounts and deadlines in consumer search. *American Economic Review*, 110(12):3748–85.
- Coey, D., Larsen, B., Sweeney, K., and Waisman, C. (2017). Ascending auctions with bidder asymmetries. *Quantitative Economics*, 8(1):181–200.
- Couts, A. (2012). Unlocking your smartphone will soon be illegal, and other dmca new rule insanity. Url:<http://www.digitaltrends.com/mobile/unlocking-your-smartphone-will-soon-be-illegal-and-other-new-dmca-rule-insanity/>. Posted October 26, 2012. Retrieved July 26, 2016.
- Couts, A. (2013). What happens if you unlock your smartphone now that it’s illegal? Url:<http://www.digitaltrends.com/mobile/what-happens-if-you-unlock-your-smartphone/>. Posted January 29, 2013. Retrieved December 6, 2019.
- David, H. A. and Nagaraja, H. N. (2003). *Order Statistic*. New York: Wiley.
- Decarolis, F. (2017). Comparing public procurement auctions. *International Economic Review*, forthcoming.
- Evdokimov, K. and White, H. (2012). Some extensions of a lemma of kotlarski. *Econometric Theory*, 28(4):925–923.

- Fan, Y. and Yang, C. (2020). Competition, product proliferation and welfare: A study of the us smartphone market. *American Economic Journal: Microeconomics*, 12(2):99–134.
- Farrell, J. and Klemperer, P. (2007). Coordination and lock-in: Competition with switching costs and network effects. *Handbook of Industrial Organization*, 3:1967–2072.
- Federal Register (2012). Exemption to prohibition on circumvention of copyright protection systems for access control technologies. *Federal Register*, 77(208).
Url:<http://copyright.gov/fedreg/2012/77fr65260.pdf>.
- Freyberger, J. and Larsen, B. (2019). Identification in ascending auctions, with an application to digital rights management. SIEPR Working paper 20-012.
- Gallant, A. and Nychka, D. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica*, pages 363–390.
- Grether, D. M., Porter, D., and Shum, M. (2016). Public vs. secret reserve prices in auctions: Evidence from combining field and lab experiments. *Secret Reserve Prices in Auctions: Evidence from Combining Field and Lab Experiments (April 18, 2016)*.
- Haile, P. A., Hong, H., and Shum, M. (2003). Nonparametric tests for common values at first-price sealed-bid auctions. NBER Working Paper 10105.
- Haile, P. A. and Tamer, E. (2003). Inference with an incomplete model of english auctions. *Journal of Political Economy*, 111(1):1–51.
- Harbaugh, R. and Khemka, R. (2010). Does copyright enforcement encourage piracy? *Journal of Industrial Economics*, 58(2):306–323.
- Hendricks, K. and Sorensen, A. (2018). Dynamics and efficiency in decentralized online auction markets. NBER Working Paper 25002.
- Hernández, C., Quint, D., and Turansick, C. (2020). Estimation in english auctions with unobserved heterogeneity. *The RAND Journal of Economics*, 51(3):868–904.
- Hickman, B. R., Hubbard, T. P., and Paarsch, H. J. (2016). Identification and estimation of a bidding model for electronic auctions. *Quantitative Economics*, forthcoming.

- Hortacsu, A. and Nielsen, E. R. (2010). Commentary: Do bids equal values on ebay? *Marketing Science*, 29(6):994–997.
- Hu, Y., McAdams, D., and Shum, M. (2013). Identification of first-price auctions with non-separable unobserved heterogeneity. *Journal of Econometrics*, 174(2):186–193.
- Judd, K. (1998). *Numerical Methods in Economics*. Cambridge: MIT Press.
- Kehoe, P., Larsen, B., and Pastorino, E. (2018). Dynamic competition in the era of big data. Working paper, Stanford University.
- Khan, S. and Tamer, E. (2010). Irregular identification, support conditions, and inverse weight estimation. *Econometrica*, 78(6):2021–2042.
- Khanna, D. (2013). The most ridiculous law of 2013 (so far): It is now a crime to unlock your smartphone. Url:<https://www.theatlantic.com/business/archive/2013/01/the-most-ridiculous-law-of-2013-so-far-it-is-now-a-crime-to-unlock-your-smartphone/272552/>. Posted January 27, 2013. Retrieved May 30, 2017.
- Kim, K. i. and Lee, J. (2014). Nonparametric estimation and testing of the symmetric ipv framework with unknown number of bidders. Working paper, Michigan State University.
- Kotlarski, I. (1967). On characterizing the gamma and normal distribution. *Pacific Journal of Mathematics*, 20(1):69–76.
- Krasnokutskaya, E. (2011). Identification and estimation of auction models with unobserved heterogeneity. *Review of Economic Studies*, 78(1):293–327.
- Krasnokutskaya, E. and Seim, K. (2011). Bid preference programs and participation in highway procurement auctions. *American Economic Review*, 101(6):2653–2686.
- Larsen, B. (2019). The efficiency of real-world bargaining: Evidence from wholesale used-auto auctions. NBER Working Paper 20431. Forthcoming, *Review of Economic Studies*.
- Li, T., Perrigne, I., and Vuong, Q. (2000). Conditionally independent private information in ocs wildcat auctions. *Journal of Econometrics*, 98(1):129–161.
- Li, T. and Vuong, Q. (1998). Nonparametric estimation of the measurement error model using multiple indicators. *Journal of Multivariate Analysis*, 65(2):139–165.

- Liu, Q., Safavi-Naini, R., and Sheppard, N. P. (2003). Digital rights management for content distribution. In *Proceedings of the Australasian information security workshop conference on ACSW frontiers 2003-Volume 21*, pages 49–58. Australian Computer Society, Inc.
- Liu, X. and Shao, Y. (2003). Asymptotics for likelihood ratio tests under loss of identifiability. *The Annals of Statistics*, 31(3):807–832.
- Luo, H. and Mortimer, J. H. (2016). Copyright enforcement: Evidence from two field experiments. NBER Working Paper 22082.
- Luo, Y. and Xiao, R. (2019). Identification of auction models using order statistics. Working paper, University of Toronto.
- Maicas, J. P., Polo, Y., and Sese, F. J. (2009). Reducing the level of switching costs in mobile communications: The case of mobile number portability. *Telecommunications Policy*, 33(9):544–554.
- Mbakop, E. (2017). Identification of auctions with incomplete bid data in the presence of unobserved heterogeneity. Working paper, Northwestern University.
- Murphy, S. A. and Van Der Vaart, A. W. (2000). On profile likelihood. *Journal of the American Statistical Association*, 95(450):449–465.
- Nakamura, A. (2010). Estimating switching costs involved in changing mobile phone carriers in japan: Evaluation of lock-in factors related to japan’s sim card locks. *Telecommunications Policy*, 34(11):736–746.
- Park, Y. and Koo, Y. (2016). An empirical analysis of switching cost in the smart-phone market in south korea. *Telecommunications Policy*, 40(4):307–318.
- Platt, B. C. (2017). Inferring ascending auction participation from observed bidders. *International Journal of Industrial Organization*, 54:65–88.
- Roberts, J. W. (2013). Unobserved heterogeneity and reserve prices in auctions. *RAND Journal of Economics*, 44(4):712–732.
- Seifert, S. and Hüttel, S. (2020). Common values and unobserved heterogeneity in farmland auctions in germany.
- Shen, X. (1997). On methods of sieves and penalization. *The Annals of Statistics*, 25(6):2555–2591.

- Shen, X. and Shi, J. (2005). Sieve likelihood ratio inference on general parameter space. *Science in China Series A: Mathematics*, 48(1):67–78.
- Sinkinson, M. (2014). Pricing and entry incentives with exclusive contracts: Evidence from smartphones. Working paper, University of Pennsylvania.
- Song, U. (2004). Nonparametric estimation of an ebay auction model with an unknown number of bidders. Working paper, University of British Columbia.
- Stallman, R. (1997). The right to read. *Communications of the ACM*, 40(2):85–87.
- Tallberg, M., Hämmäinen, H., Töyli, J., Kamppari, S., and Kivi, A. (2007). Impacts of handset bundling on mobile data usage: The case of finland. *Telecommunications Policy*, 31(10):648–659.
- Van Camp, J. (2013). Stick it to the man! phone unlocking skyrockets 70 percent since becoming illegal. Url:<http://www.digitaltrends.com/mobile/phone-unlocking-skyrockets-since-ban/>. Posted August 13, 2013. Retrieved September 1, 2016.
- Velazco, C. (2013). Unlocking your phone is now illegal, but what does that mean for you? Url:<https://techcrunch.com/2013/01/26/unlocking-your-phone-is-now-illegal-but-what-does-that-mean-for-you/>. Posted January 26, 2013. Retrieved May 30, 2017.
- Von Lohmann, F. (2004). Measuring the digital millennium against the darknet: Implications for the regulation of technological protection measures. *Loyola of Los Angeles Entertainment Law Review*, 24:635.
- Walker, J. (2003). The digital imprimatur: How big brother and big media can put the internet genie back in the bottle. *Knowledge, Technology, and Policy*, 16(3):24–77.
- Wyatt, E. (2013). Fcc backs consumers in unlocking of cellphones. *New York Times*, page B1. March 5, 2013.
- Yang, C. (2020). Vertical structure and innovation: A study of the soc and smart-phone industries. *The RAND Journal of Economics*, 51(3):739–785.
- Zhu, T., Liu, H., and Chintagunta, P. K. (2015). Wireless carriers’ exclusive handset arrangements: An empirical look at the iphone. *Customer Needs and Solutions*, 2(2):177–190.

TABLE 1. Descriptive Statistics

	Mean	Std. Dev.	Min	Max
Unlocked	.114	.318	0	1
AT&T	.55	.498	0	1
Verizon	.336	.472	0	1
iPhone Model 4	.753	.431	0	1
8GB Memory	.0809	.273	0	1
16GB Memory	.698	.459	0	1
32GB Memory	.221	.415	0	1
Black Phone	.771	.42	0	1
Post Ban	.33	.47	0	1
Reserve Price	229	77.2	80	475
Second Highest Bid	247	71.5	100	440
Third Highest Bid	229	70.8	90	420
Minutes to End, Highest Bid	127	725	0	13,925
Minutes to End, Second Highest Bid	158	752	0	14,261
Minutes to End, Third Highest Bid	335	1,149	0	14,379
# Observed Bidders	12	4.01	3	30
# Bids	30.4	14.4	0	130
# Photos	4.05	3	0	12
Shipping Fee	5.67	5.35	0	20
Number of Listings	12,872			

Notes: Table displays descriptive statistics for eBay used iPhone data.

TABLE 2. Descriptive Statistics Differences and Difference-in-Differences

	Locked Phones			Unlocked Phones			Diff-in-diff
	Pre	Post	Diff	Pre	Post	Diff	
Reserve Price	226.7 (74.0)	214.4 (74.8)	-12.2 (1.5)	278.9 (80.8)	275.1 (78.7)	-3.8 (4.4)	8.5 (4.6)
Second Highest Bid	247.4 (68.0)	228.4 (73.4)	-19.0 (1.4)	289.2 (68.2)	285.1 (61.6)	-4.1 (3.5)	14.9 (3.8)
# Observed Bidders	11.8 (3.9)	11.9 (4.0)	0.1 (0.1)	12.8 (4.0)	13.6 (4.6)	0.7 (0.2)	0.7 (0.3)
# Bids	29.2 (13.8)	31.2 (14.6)	2.0 (0.3)	33.7 (14.7)	37.1 (16.4)	3.4 (0.9)	1.3 (0.9)
Total # Sellers	5,604	2,546	-3,058	656	341	-315	2,743
Total # Listings	7,661	3,745	-3,916	966	500	-466	3,450

Notes: Table displays means and standard deviations (in parentheses) of several variables—reserve price, second-highest bid, number of observed bidders, and number of observed bids—broken down by auctions for locked vs. unlocked phones and by time period (pre vs. post the unlocking ban). Columns labeled “Diff” show the difference pre vs. post, along with standard errors in parentheses. Column labeled “Diff-in-diff” shows the difference between the “Diff” columns, along with standard errors in parentheses. Final two rows show the total number of sellers and listing in each subsample, along with differences and difference-in-differences. Bolded entries indicate t-statistics of magnitude greater than 1.96.

TABLE 3. Regression Results: Number of Bidders and Log(Reserve+1)

	Number of Observed Bidders			Log(Reserve+1)		
	(1)	(2)	(3)	(4)	(5)	(6)
Log(Photos+1)	-0.0130 (0.0538)		0.00948 (0.0526)	0.0119** (0.00450)		0.0117** (0.00357)
iPhone 4, 8 GB, W		-0.176 (0.236)	-0.176 (0.236)		0.0100 (0.0184)	0.00909 (0.0184)
iPhone 4, 16 GB, B		0.709*** (0.171)	0.708*** (0.171)		0.0805*** (0.0125)	0.0793*** (0.0125)
iPhone 4, 16 GB, W		-0.114 (0.209)	-0.115 (0.209)		0.0970*** (0.0160)	0.0953*** (0.0161)
iPhone 4, 32 GB, B		-0.153 (0.184)	-0.154 (0.184)		0.219*** (0.0135)	0.218*** (0.0135)
iPhone 4, 32 GB, W		-0.955** (0.321)	-0.957** (0.321)		0.213*** (0.0269)	0.210*** (0.0269)
iPhone 4S, 16 GB, B		2.139*** (0.204)	2.138*** (0.204)		0.508*** (0.0142)	0.507*** (0.0142)
iPhone 4S, 16 GB, W		2.453*** (0.208)	2.452*** (0.208)		0.542*** (0.0140)	0.541*** (0.0140)
iPhone 4S, 32 GB, B		1.826*** (0.251)	1.825*** (0.251)		0.610*** (0.0170)	0.609*** (0.0170)
iPhone 4S, 32 GB, W		1.258*** (0.264)	1.257*** (0.264)		0.641*** (0.0157)	0.639*** (0.0157)
Constant	12.00*** (0.0843)	11.19*** (0.163)	11.17*** (0.178)	5.387*** (0.00706)	5.190*** (0.0121)	5.175*** (0.0129)
R^2	0.00000458	0.0461	0.0461	0.000559	0.370	0.370

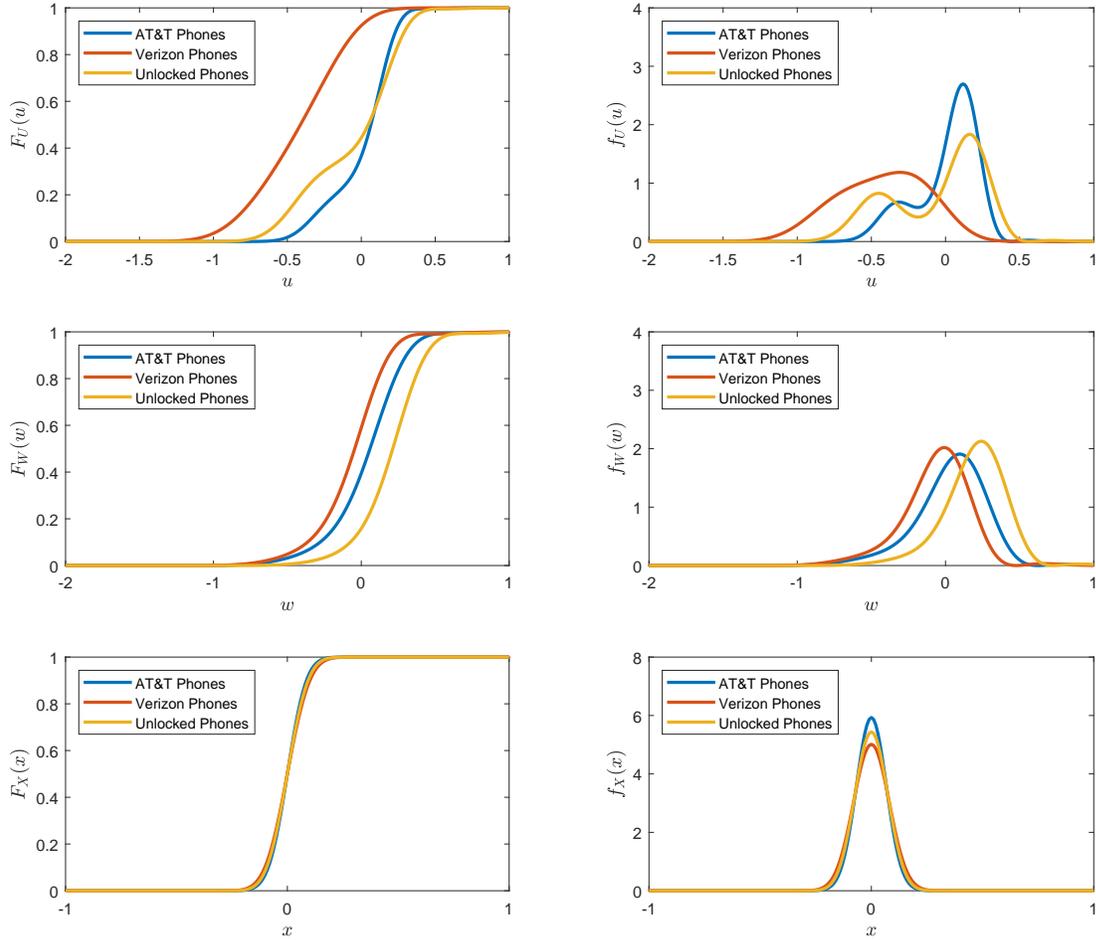
Notes: Columns 1–3 display results from regressing the number of observed bidders in a given auction on auction-level observable characteristics, including the number of photos and categorical dummies for the type of iPhone. In columns 4–6 the dependent variable is log(reserve+1). “W”=White, “B”=Black. * = $p \leq 0.10$, ** = $p \leq 0.05$, *** = $p \leq 0.01$.

TABLE 4. Model Estimates

A. Semiparametric Model, Full Sample											
	AT&T Phones			Verizon Phones			Unlocked Phones			Diff-in-diff	
	Pre	Post	Diff	Pre	Post	Diff	Pre	Post	Diff	AT&T	Verizon
E(U)	0.015 (0.014)	-0.008 (0.027)	-0.024 (0.031)	-0.333 (0.027)	-0.372 (0.027)	-0.039 (0.038)	0.005 (0.039)	-0.164 (0.079)	-0.169 (0.088)	-0.145 (0.093)	-0.130 (0.096)
Std(U)	0.153 (0.006)	0.177 (0.011)	0.023 (0.012)	0.240 (0.010)	0.220 (0.011)	-0.021 (0.015)	0.200 (0.014)	0.257 (0.022)	0.057 (0.026)	0.034 (0.029)	0.078 (0.030)
E(W)	0.025 (0.003)	0.030 (0.006)	0.005 (0.007)	-0.070 (0.005)	-0.144 (0.006)	-0.074 (0.008)	0.183 (0.007)	0.168 (0.012)	-0.015 (0.015)	-0.020 (0.016)	0.058 (0.016)
Std(W)	0.241 (0.003)	0.243 (0.006)	0.002 (0.007)	0.229 (0.005)	0.238 (0.005)	0.008 (0.007)	0.222 (0.009)	0.261 (0.016)	0.039 (0.018)	0.037 (0.019)	0.031 (0.019)
Std(X)	0.066 (0.002)	0.064 (0.006)	-0.002 (0.006)	0.072 (0.007)	0.083 (0.007)	0.011 (0.010)	0.065 (0.008)	0.088 (0.015)	0.022 (0.017)	0.024 (0.018)	0.011 (0.020)
N	5,251	1,829		2,410	1,916		966	500			
B. Semiparametric Model, Late-Bidding Sample											
E(U)	0.026 (0.013)	0.026 (0.026)	0.000 (0.029)	-0.435 (0.051)	-0.471 (0.054)	-0.036 (0.074)	0.045 (0.058)	-0.176 (0.141)	-0.221 (0.152)	-0.221 (0.155)	-0.185 (0.169)
Std(U)	0.128 (0.005)	0.135 (0.009)	0.007 (0.010)	0.241 (0.014)	0.227 (0.017)	-0.014 (0.023)	0.147 (0.020)	0.195 (0.028)	0.048 (0.035)	0.041 (0.036)	0.062 (0.041)
E(W)	0.005 (0.005)	-0.001 (0.009)	-0.007 (0.010)	-0.105 (0.007)	-0.173 (0.008)	-0.068 (0.011)	0.147 (0.013)	0.094 (0.021)	-0.053 (0.025)	-0.046 (0.027)	0.015 (0.027)
Std(W)	0.237 (0.004)	0.250 (0.008)	0.013 (0.009)	0.228 (0.006)	0.238 (0.007)	0.010 (0.009)	0.230 (0.014)	0.256 (0.021)	0.026 (0.025)	0.013 (0.027)	0.016 (0.027)
Std(X)	0.054 (0.003)	0.071 (0.005)	0.016 (0.006)	0.077 (0.008)	0.082 (0.008)	0.005 (0.011)	0.069 (0.008)	0.097 (0.012)	0.028 (0.014)	0.012 (0.015)	0.023 (0.018)
N	2,745	869		1,135	1,006		353	166			
C. Nonparametric Model, Full Sample											
E(U)	0.010 (0.000)	-0.067 (0.015)	-0.077 (0.015)	-0.429 (0.037)	-0.387 (0.026)	0.042 (0.045)	-0.055 (0.019)	-0.567 (0.053)	-0.512 (0.056)	-0.435 (0.058)	-0.554 (0.072)
Std(U)	0.213 (0.000)	0.285 (0.002)	0.072 (0.002)	0.309 (0.014)	0.233 (0.014)	-0.077 (0.020)	0.320 (0.001)	0.352 (0.037)	0.032 (0.037)	-0.040 (0.037)	0.108 (0.042)
E(W)	0.033 (0.002)	0.038 (0.004)	0.005 (0.005)	-0.068 (0.005)	-0.149 (0.005)	-0.081 (0.007)	0.194 (0.006)	0.197 (0.011)	0.003 (0.013)	-0.003 (0.014)	0.083 (0.015)
Std(W)	0.243 (0.003)	0.256 (0.006)	0.012 (0.006)	0.233 (0.005)	0.239 (0.005)	0.006 (0.007)	0.214 (0.006)	0.275 (0.011)	0.061 (0.012)	0.048 (0.014)	0.055 (0.014)
Std(X)	0.067 (0.002)	0.060 (0.004)	-0.007 (0.004)	0.079 (0.005)	0.086 (0.005)	0.007 (0.007)	0.073 (0.004)	0.078 (0.007)	0.006 (0.008)	0.013 (0.009)	-0.002 (0.010)
N	5,251	1,829		2,410	1,916		966	500			
D. Nonparametric Model, Late-Bidding Sample											
E(U)	-0.030 (0.005)	-0.118 (0.010)	-0.087 (0.012)	-0.604 (0.082)	-0.495 (0.113)	0.110 (0.139)	0.069 (0.023)	-0.503 (0.026)	-0.572 (0.035)	-0.484 (0.037)	-0.681 (0.143)
Std(U)	0.163 (0.002)	0.204 (0.004)	0.041 (0.004)	0.294 (0.013)	0.242 (0.062)	-0.053 (0.063)	0.147 (0.044)	0.177 (0.014)	0.030 (0.046)	-0.011 (0.046)	0.082 (0.078)
E(W)	0.011 (0.002)	0.008 (0.005)	-0.004 (0.005)	-0.104 (0.007)	-0.176 (0.008)	-0.072 (0.010)	0.157 (0.008)	0.111 (0.016)	-0.046 (0.017)	-0.042 (0.018)	0.026 (0.020)
Std(W)	0.241 (0.005)	0.261 (0.007)	0.021 (0.009)	0.241 (0.006)	0.236 (0.006)	-0.004 (0.009)	0.223 (0.009)	0.260 (0.017)	0.037 (0.019)	0.016 (0.021)	0.041 (0.021)
Std(X)	0.055 (0.003)	0.068 (0.003)	0.012 (0.004)	0.066 (0.008)	0.084 (0.007)	0.018 (0.011)	0.066 (0.007)	0.097 (0.008)	0.031 (0.011)	0.018 (0.012)	0.013 (0.015)
N	2,745	869		1,135	1,006		353	166			

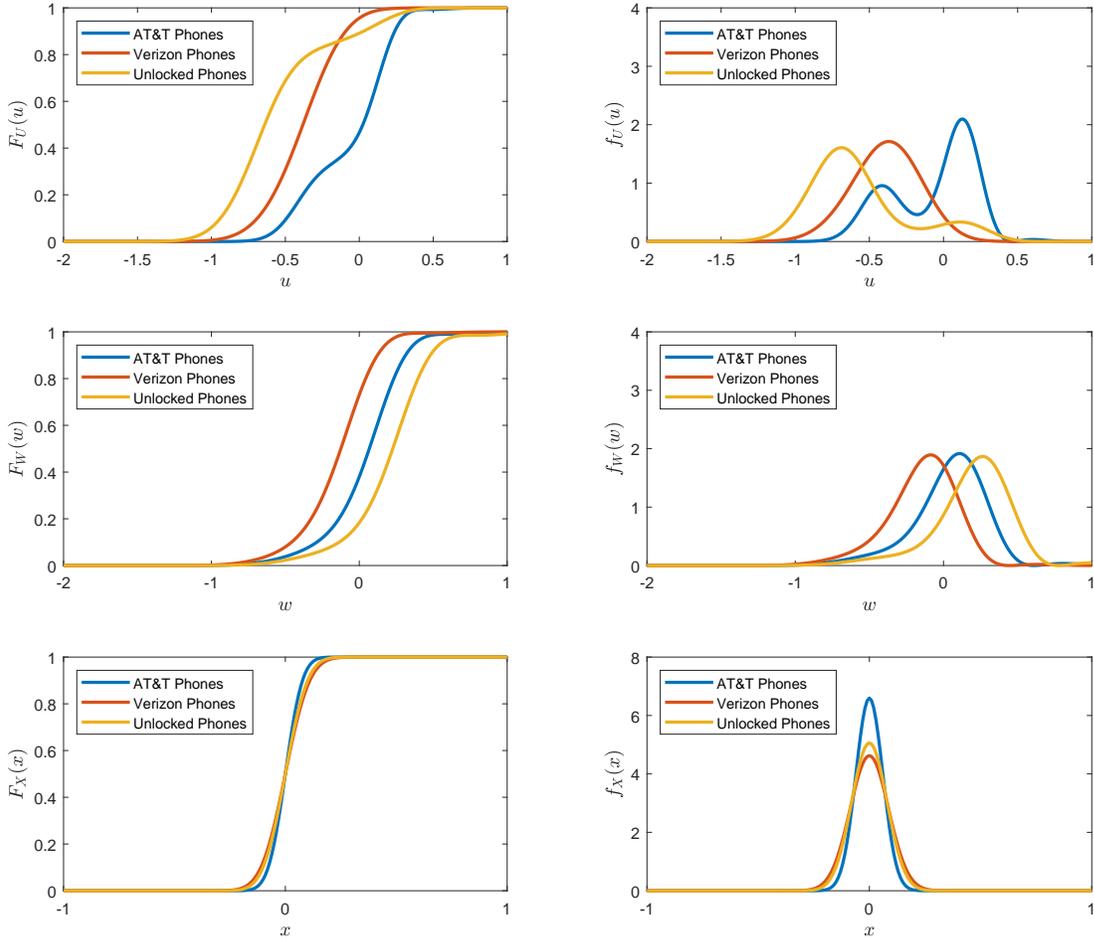
Notes: Pre and Post columns display model estimates of means and standard deviations before and after the unlocking ban. Diff columns display Post estimate minus Pre estimate. This analysis is performed separately for AT&T, Verizon, and Unlocked phones. AT&T Diff-in-diff column displays the difference-in-differences between unlocked and AT&T phones, and analogously for Verizon column. Standard errors (in parentheses) for each quantity are estimated using the delta method, treating the model as parametric. Bolded entries indicate t-statistics greater than 1.96 in absolute value.

FIGURE 1. Pre Unlocking Ban: Distribution Functions and Densities for AT&T, Verizon, and Unlocked Phones



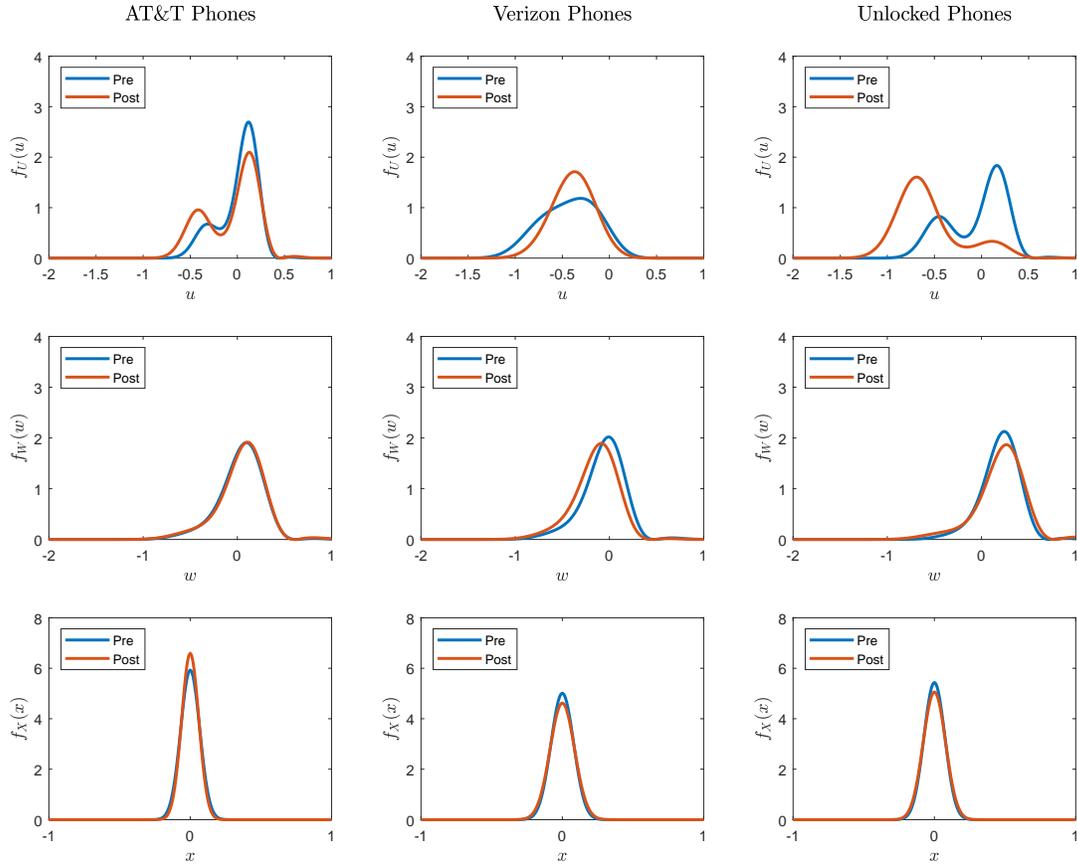
Notes: Each plot shows nonparametric estimates for AT&T, Verizon, and unlocked phones in the pre period. Panels on the left show distribution functions and on the right, densities, for the non-common component of buyer valuations F_U (top row) and reserve prices F_W (middle row), and for the unobserved heterogeneity F_X (bottom row). Units on the horizontal axes are log points, after homogenization (i.e. subtracting off observable auction-level heterogeneity).

FIGURE 2. Post Unlocking Ban: Distribution Functions and Densities for AT&T, Verizon, and Unlocked Phones



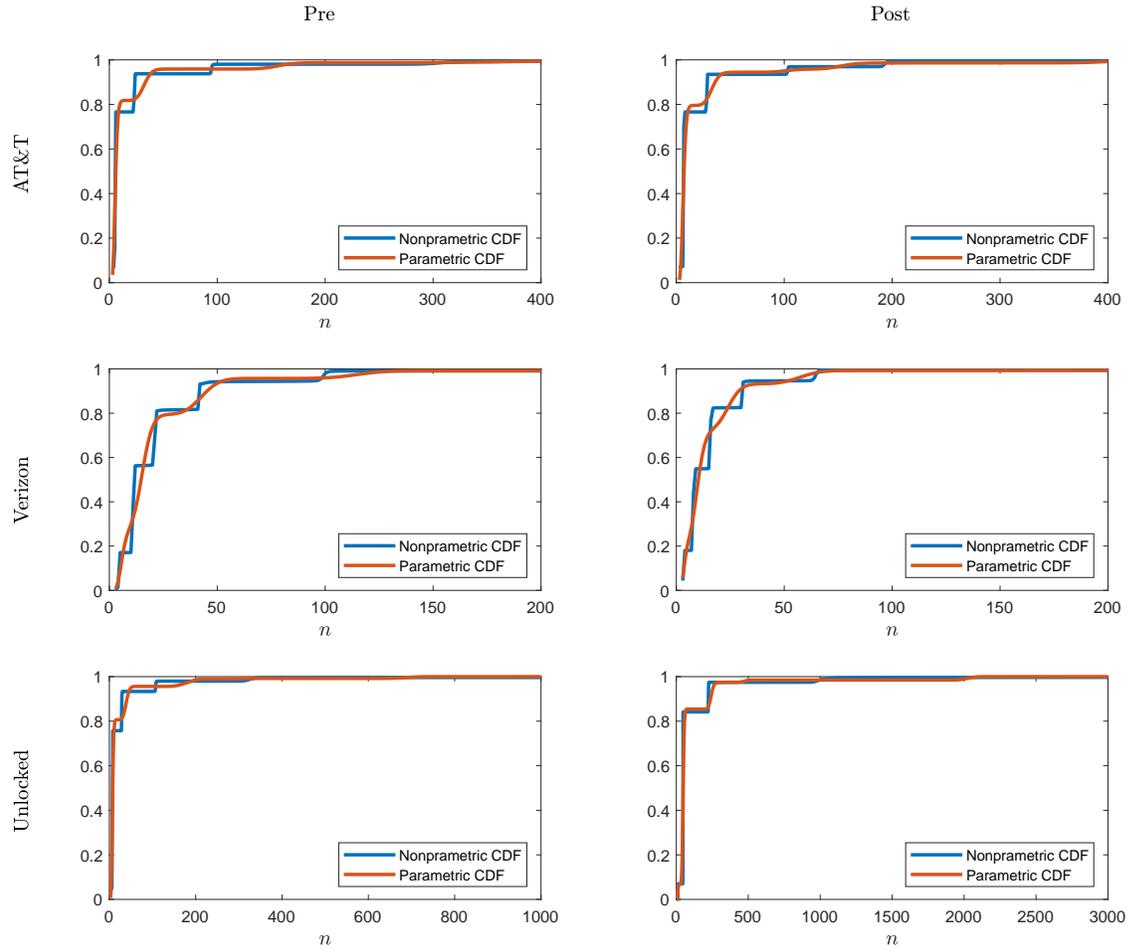
Notes: Each plot shows nonparametric estimates for AT&T, Verizon, and unlocked phones in the post period. Panels on the left show distribution functions and on the right, densities, for the non-common component of buyer valuations F_U (top row) and reserve prices F_W (middle row), and for the unobserved heterogeneity F_X (bottom row). Units on the horizontal axes are log points, after homogenization (i.e. subtracting off observable auction-level heterogeneity).

FIGURE 3. Estimates Before and After Unlocking Ban



Notes: Each plot shows the nonparametric density estimates before and after the regulatory change. Panels on the left show estimates for AT&T phones, panels in the middle for Verizon phones, and panels on the right for unlocked phones. The densities belong to the non-common component of buyer valuations F_U (top row) and the reserve prices F_W (middle row), and the unobserved heterogeneity F_X (bottom row). Units on the horizontal axes are log points, after homogenization (i.e. subtracting off observable auction-level heterogeneity).

FIGURE 4. Estimates for Distribution of Number of Bidders Before and After Unlocking Ban



Notes: Each plot shows the a parametric (Poisson mixture) and nonparametric estimate of the CDF of the number of bidders. Estimates from the pre period are shown on the left and from the post period on the right for AT&T (top row), Verizon (middle row), and unlocked phones (bottom row). Note that the range on the horizontal axis is different for AT&T, Verizon, and unlocked phones.

APPENDIX A. PROOFS

Proof of Theorem 1. First, write $B^j = X + U^j$ and $R = X + W$. By Lemma 1 of Evdokimov and White (2012), ϕ_X is identified. Moreover, notice that by independence of X and (U^j, U^k, W) we get

$$\phi_{B^j, B^k, R}(t_j, t_k, t) = \phi_X(t_j + t_k + t)\phi_{U^j, U^k, W}(t_j, t_k, t)$$

for all $(t_j, t_k, t) \in \mathbb{R}^3$. Therefore, for all $(t_j, t_k, t) \in \mathbb{R}^3$ such that $\phi_X(t_j + t_k + t) \neq 0$,

$$\phi_{U^j, U^k, W}(t_j, t_k, t) = \frac{\phi_{B^j, B^k, R}(t_j, t_k, t)}{\phi_X(t_j + t_k + t)}.$$

It follows that $\phi_{U^j, U^k, W}(t_j, t_k, t)$ is identified for all $(t_j, t_k, t) \in \mathbb{R}^3$ such that $\phi_X(t_j + t_k + t) \neq 0$. Since the zeros of ϕ_X are isolated and since $\phi_{U^j, U^k, W}(t_j, t_k, t)$ is continuous, $\phi_{U^j, U^k, W}(t_j, t_k, t)$ is identified for all $(t_j, t_k, t) \in \mathbb{R}^3$. Identification of the characteristic function is equivalent to identification of the density $f_{U^j, U^k, W}$.

Since we know the joint distribution of U^j and U^k , we can use arguments as in Song (2004) to identify the distribution of valuations. Specifically, Song (2004) demonstrated that the density of U^j conditional on a realization of U^k does not depend on the realization of N :

$$\begin{aligned} f_{U^j|U^k}(u_j|u_k) &= \frac{f_{U^j, U^k}(u_j, u_k)}{f_{U^k}(u_k)} \\ &= \frac{(k-1)!(F_U(u_j) - F_U(u_k))^{k-j-1}(1 - F_U(u_j))^{j-1}f_U(u_j)}{(k-j-1)!(j-1)!(1 - F_U(u_k))^{k-1}}. \end{aligned}$$

Letting u_k converge to the lower bound of the support of U identifies

$$\frac{(k-1)F_U(u_j)^{k-j-1}(1 - F_U(u_j))^{j-1}f_U(u_j)}{(k-j-1)!(j-1)!},$$

which is the density of the $(j-k)$ 'th order statistic from a sample of size $(j-1)$ with distribution function F_U . Hence, from Theorem 1 in Athey and Haile (2002), F_U is identified.

Finally, note that if $N \in \{\underline{n}, \dots, \bar{n}\}$ with $\underline{n} \geq k$ and $\bar{n} < \infty$, then

$$\begin{aligned} F_{U^k}(u) &= \sum_{n=\underline{n}}^{\bar{n}} \Pr(N = n) \left[\sum_{i=n-k+1}^n \binom{n}{i} F_U(u)^i (1 - F_U(u))^{n-i} \right] \\ &= \sum_{n=\underline{n}}^{\bar{n}} \Pr(N = n) \left[\sum_{i=n-k+1}^n \binom{n}{i} F_U(u)^i \sum_{r=1}^{n-i} \binom{n-i}{r} (-1)^r F_U(u)^r \right] \\ &= \sum_{n=\underline{n}}^{\bar{n}} \Pr(N = n) \left[\sum_{i=n-k+1}^n \sum_{r=1}^{n-i} \binom{n-i}{r} \binom{n}{i} (-1)^r F_U(u)^{i+r} \right] \end{aligned}$$

which is a polynomial in F_U . Also notice that $F_U(u)$ is identified for all u and its range is $(0, 1)$ because it is continuous. The coefficient in front of $F_U(u)^{\bar{n}}$, which is the highest order polynomial, is therefore identified and is

$$\Pr(N = \bar{n}) \sum_{i=\bar{n}-k+1}^{\bar{n}} \binom{\bar{n}}{i} (-1)^{\bar{n}-i}.$$

Hence, $\Pr(N = \bar{n})$ is identified as long as $\sum_{i=\bar{n}-k+1}^{\bar{n}} \binom{\bar{n}}{i} (-1)^{\bar{n}-i} \neq 0$. Indeed,

$$\sum_{i=\bar{n}-k+1}^{\bar{n}} \binom{\bar{n}}{i} (-1)^{\bar{n}-i} = \sum_{j=1}^{k-1} \binom{\bar{n}}{\bar{n}-j} (-1)^j = \sum_{j=1}^{k-1} \binom{\bar{n}}{j} (-1)^j = (-1)^{k-1} \binom{\bar{n}-1}{k-1} \neq 0.$$

Given identification of $\Pr(N = \bar{n})$, we know

$$F_{U^k}(u) - \Pr(N = \bar{n}) \left[\sum_{i=\bar{n}-k+1}^{\bar{n}} \binom{\bar{n}}{i} F_U(u)^i (1 - F_U(u))^{\bar{n}-i} \right],$$

which equals

$$\sum_{n=\underline{n}}^{\bar{n}-1} \Pr(N = n) \left[\sum_{i=n-k+1}^n \binom{n}{i} F_U(u)^i (1 - F_U(u))^{n-i} \right].$$

Analogous arguments as above now show identification of $\Pr(N = \bar{n} - 1)$. Repeating these steps yield identification of $\Pr(N = n)$ for all $n \in \{\underline{n}, \dots, \bar{n}\}$.

Proof of Theorem 2. Since $\bar{B} \geq \bar{R}$, \bar{B} is identified from the largest possible bid. Notice that if $B^j = \bar{B}$, then $U^j = \bar{U}$ and $X = \bar{X}$. We can now observe for all

$$r \in [\underline{R}, \overline{R}] = [\underline{X} + \underline{W}, \overline{X} + \overline{W}]$$

$$\begin{aligned} P(R \leq r \mid B^j = \overline{B}) &= P(R \leq r \mid X = \overline{X}, U^j = \overline{U}) \\ &= P(W \leq r - \overline{X}). \end{aligned}$$

It follows that the distribution of $\tilde{W} = W + \overline{X}$ is identified. But since $E[\tilde{W}] = E[W] + \overline{X} = E[R] + \overline{X}$ and since $E[R]$ is identified, it follows that \overline{X} and hence the distribution of W is identified. By independence of X and W , $\phi_R(t) = \phi_X(t)\phi_W(t)$, and for all t with $\phi_W(t) \neq 0$,

$$\phi_X(t) = \frac{\phi_R(t)}{\phi_W(t)}.$$

Since $\phi_R(t)$ and $\phi_W(t)$ are identified and since $\phi_W(t)$ has only isolated zeros, $\phi_X(t)$ is identified. Thus, F_X is identified.

Moreover,

$$\begin{aligned} P(B^j \geq b_j, B^k \geq b_k \mid R = \underline{R}) &= P(B^j \geq b_j, B^k \geq b_k \mid X + W = \underline{X} + \underline{W}) \\ &= P(B^j \geq b_j, B^k \geq b_k \mid X = \underline{X}, W = \underline{W}) \\ &= P(U^j \geq b_j - \underline{X}, U^k \geq b_k - \underline{X} \mid X = \underline{X}, W = \underline{W}) \\ &= P(U^j \geq b_j - \underline{X}, U^k \geq b_k - \underline{X}) \end{aligned}$$

Now suppose that $\underline{B} \geq \underline{R}$ implies that $\underline{U} \geq \underline{W}$. Given $X = \underline{X}$, the support of the observed B^j is $[\underline{X} + \underline{W}, \underline{X} + \overline{U}]$. Let $u_j, u_k \in [\underline{U}, \overline{U}]$ and let $b_j = u_j + \underline{X}$ and $b_k = u_k + \underline{X}$, which are on the support of B^j and B^k . It follows from the previous equation that

$$P(U^j \geq u_j, U^k \geq u_k) = P(B^j \geq u_j + \underline{X}, B^k \geq u_k + \underline{X} \mid R = \underline{R}).$$

for all $u_j, u_k \in [\underline{U}, \overline{U}]$. Since the right hand side is identified, F_{U^j, U^k} is identified. Just as in the proof of Theorem 1, identification of F_{U^j, U^k} yields identification of F_U and $P(N = n)$.

If instead $\underline{B} < \underline{R}$, we can only identify $P(U^j \geq u_j, U^k \geq u_k)$ for all $u_j \geq u_k \geq \underline{W} > \underline{U}$. Analogous arguments identify $P(U^j \geq u_j)$ and $P(U^k \geq u_k)$ for all $u_j, u_k \geq \underline{W} >$

\underline{U} . Now since

$$F_{U^j|U^k}(u_j | u_k) = 1 - \int_{u_j}^{\bar{U}} \frac{f_{U^j,U^k}(z, u_k)}{f_{U^k}(u_k)} dz,$$

$F_{U^j|U^k}(u_j | u_k)$ for all $u_j \geq u_k \geq \underline{W}$ is identified as well. Therefore, by the arguments from Song (2004), knowing $F_{U^j|U^k}(u_j | u_k)$ implies knowledge of

$$(F_U(u_j) - F_U(u_k))/(1 - F_U(u_k))$$

for all $u_j \geq u_k \geq \underline{W}$. In other words, we can identify $P(U \leq u_j | U \geq u_k)$ for all $u_j, u_k \geq \underline{W}$.

APPENDIX B. LIKELIHOOD DERIVATION

Here we derive the expression for p_1 , p_2 , and p_3 given in Section 3.2. For the first part write

$$\begin{aligned} P(R \leq r, D_1 = 1) &= P(R \leq r, B^k \leq B^j < R) \\ &= P(R \leq r, B^j < R) \\ &= \int_{-\infty}^{\infty} P(W \leq r - x, U^j \leq W) f_X(x) dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(W \leq r - x, u \leq W | U^j = u_j) f_{U^j}(u_j) f_X(x) du_j dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{r-x} P(u_j \leq W \leq r - x) f_{U^j}(u_j) f_X(x) du_j dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{r-x} (P(W \leq r - x) - P(W \leq u_j)) f_{U^j}(u_j) f_X(x) du_j dx \end{aligned}$$

Taking the derivative with respect to r (using Leibniz's rule) yields

$$\begin{aligned} p_1(r) &= \int_{-\infty}^{\infty} \int_{-\infty}^{r-x} f_W(r - x) f_{U^j}(u_j) f_X(x) du_j dx \\ &= \int_{-\infty}^{\infty} F_{U^j}(r - x) f_W(r - x) f_X(x) dx. \end{aligned}$$

If the number of bidders is unknown, we need expressions in terms of the distributions of $U^j | U^k$ and U^k and we use

$$F_{U^j}(u_j) = \int_{-\infty}^{-\infty} F_{U^j|U^k}(u_j | u_k) f_{U^k}(u_k) du_k.$$

Specifically, when B^j is the second highest and B^k is the third highest bid, then

$$f_{U^j|U^k}(u_j, | u_k) = \frac{2(1 - F_U(u_j))f_U(u_j)}{(1 - F_U(u_k))^2}.$$

Thus,

$$\begin{aligned} F_{U^j|U^k}(u_j, | u_k) &= \int_{u_k}^{u_j} \frac{2(1 - F_U(z))f_U(z)}{(1 - F_U(u_k))^2} dz \\ &= \frac{(2F_U(u_j) - F_U(u_j)^2) - (2F_U(u_k) - F_U(u_k)^2)}{(1 - F_U(u_k))^2} \mathbf{1}(u_j \geq u_k). \end{aligned}$$

Similarly,

$$\begin{aligned} P(B^j \leq b_j, R \leq r, D_2 = 1) &= P(B^j \leq b_j, R \leq r, B^j \geq R > B^k) \\ &= \int_{-\infty}^{\infty} P(U^j \leq b_j - x, W \leq r - x, U^j \geq W > U^k) f_X(x) dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(U^j \leq b_j - x, w \leq r - x, U^j \geq w > U^k | W = w) f_W(w) f_X(x) dw dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\min\{b_j - x, r - x\}} P(w \leq U^j \leq b_j - x, U^k < w) f_W(w) f_X(x) dw dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\min\{b_j - x, r - x\}} (P(U^j \leq b_j - x, U^k < w) - P(U^j \leq w, U^k < w)) f_W(w) f_X(x) dw dx. \end{aligned}$$

If $b_j < r$, then $P(B^j \leq b_j, R \leq r, D_2 = 1)$ does not depend on r and $p_2(r, b_j) = 0$. If

$b_j \geq r$, then

$$\begin{aligned} p_2(b, r) &= \frac{\partial}{\partial r \partial b_j} P(B^j \leq b_j, R \leq r, D_2 = 1) \\ &= \int_{-\infty}^{\infty} \frac{\partial}{\partial b_j} P(U^j \leq b_j - x, U^k < r - x) f_W(r - x) f_X(x) dx \\ &= \int_{-\infty}^{\infty} F_{U^k|U^j}(r - x | b_j - x) f_{U^j}(b_j - x) f_W(r - x) f_X(x) dx. \end{aligned}$$

Alternatively in terms of the distributions of $U^j | U^k$ and U^k write

$$\begin{aligned} \frac{\partial}{\partial b_j} P(U^j \leq b_j - x, U^k < r - x) &= \frac{\partial}{\partial b_j} \int_{-\infty}^{b_j - x} \int_{-\infty}^{r - x} f_{U^j, U^k}(u_j, u_k) du_j du_k \\ &= \int_{-\infty}^{r - x} f_{U^j, U^k}(b_j - x, u_k) du_k \\ &= \int_{-\infty}^{r - x} f_{U^j | U^k}(b_j - x | u_k) f_{U^k}(u_k) du_k. \end{aligned}$$

Then

$$p_2(r, b_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{r - x} f_{U^j | U^k}(b_j - x | u_k) f_{U^k}(u_k) du_k f_W(r - x) f_X(x) dx.$$

Finally, using the same arguments as before

$$\begin{aligned} p_3(r, b_j, b_2) &= \int_{-\infty}^{\infty} f_{U^j, U^k}(b_j - x, b_k - x) f_W(r - x) f_X(x) dx \\ &= \int_{-\infty}^{\infty} f_{U^j | U^k}(b_j - x | b_k - x) f_{U^k}(b_k - x) f_W(r - x) f_X(x) dx. \end{aligned}$$