Endogeneity in Entry Games: U.S. Cellphone Service Deployment*

Zhongjian Lin† Emory Univeristy
Xun Tang‡ Rice University
Mo Xiao§ University of Arizona

August 11, 2021

Abstract

In some Bayesian games, payoff-relevant states are influenced by unobserved heterogeneity that also directly affects strategic decisions. When ignored, such endogeneity leads to erroneous parameter inference and policy implications. We introduce a control-function approach for estimating such discrete Bayesian games, and apply the method to an entry game of deploying 4G-LTE technology by major U.S. cellphone service providers. Taking network investment as endogenous, we find that a hypothetical T-Mobile and Sprint merger would reduce 4G-LTE deployment across local markets and disproportionately decrease rural coverage. Ignoring such endogeneity would under-predict the negative impacts of the merger, therefore favoring its approval.

Keywords: Endogeneity, Discrete Bayesian Games, Control Function, Two-Step Nested Pseudo Likelihood, Entry Game, U.S. Cellphone Service.

JEL Classifications: C31; C35; C57; L13; L96.

---

*We are grateful to Andres Aradillas-Lopez, Maria Arbatskaya, Christoph Breunig, Keisuke Hirano, Krzysztof Karbownik, Ruixuan Liu, Essie Maasoumi, Sue Mialon and Joris Pinkse for useful comments and suggestions. All remaining errors are our own.

†Department of Economics, Emory University. 1602 Fishburne Dr., Atlanta, GA 30307. Email: zhongjian.lin@emory.edu

‡Department of Economics, Rice University. 6100 Main St. Houston, TX 77005. Email: xun.tang@rice.edu

§Department of Economics, Eller College of Management, the University of Arizona, Tucson, AZ 85721. Email: mxiao@eller.arizona.edu

1 Introduction

We propose a control-function approach for estimating discrete Bayesian games when observable payoff-relevant covariates are endogenous. Bayesian games provide a powerful framework for analyzing strategic interaction between individuals or firms with private information (a.k.a. types or signals) and have been studied in a wide range of applied contexts. Examples include choices of effort by students and teachers in classrooms in Todd and Wolpin (2018); choices of fitness exercises by adolescents in Jackson, Lin, and Yu (2020); location choices in video retail industry in Seim (2006); timing of commercials by radio stations in Sweeting (2009); movie release date timing game in Einav (2010); and market entry and exit of grocery stores in Grieco (2014). An important assumption required for inference in these empirical studies is that the covariates, at both the player- and the game-level, are exogenous.

When covariates in a Bayesian game are endogenous, identification and estimation require further assumptions on the joint distribution of covariates *and* private types. First, this poses challenges that are analogous to endogeneity in single-agent qualitative response models but are aggravated in settings with strategic interaction. More importantly, if such endogeneity is due to unobserved market-level or game-level heterogeneity that also influences the covariates and types of other players, then the private types are generally correlated even after conditioning on all covariates. This complicates the equilibrium characterization, as well as identification and estimation.[1]

Endogeneity in covariates is common in environments with strategic interaction. For example, consider the decisions by cellphone service providers to deploy a new generation of cellphone technology in local service markets. These providers rely on cellular network infrastructures, such as transmission facilities and switching offices, to provide cellphone services. As the technology evolved from 3G to 4G-LTE in the last decade, a provider could reconfigure and upgrade its 3G network to deliver 4G-LTE services. In addition, the spillover effects of a provider's 4G-LTE deployment in neighboring markets could reduce the deployment costs in a focal market. Therefore, a provider's 3G deployment in a local market and 4G-LTE deployment in neighboring markets are both important covariates that influence its decision to enter a local 4G-LTE

[1] To see this, consider a binary game with two players i, j, individual-specific covariates X_i, X_j, and private types u_i, u_j. Suppose that X_i, X_j and u_i, u_j are all correlated through some unknown market-level or game-level factor v_0. In this case, conditioning on (X_i, X_j) is not sufficient for attaining independence between (u_i, u_j) in general. Thus i's equilibrium belief about j's decision D_j would be a non-trivial function of its own types $E(D_j|X_i, X_j, u_i)$.

market. Endogeneity in these covariates may result from several sources. For instance, there are unreported demographic or geographic characteristics (e.g., topographic features) that affect both existing 3G deployment and the costs of upgrading facilities for 4G-LTE in a focal market. Besides, a provider's spectrum holdings are not reported in the data, but are strongly (if not perfectly) correlated between focal and neighboring markets. More broadly speaking, endogeneity is a concern in social-economic settings in which payoff-relevant states are influenced by unreported player- or game-level factors. We are not aware of any paper that allows for such flexible sources of endogeneity in empirical analysis of discrete Bayesian games and investigates the impact on inference and policy implications when endogeneity is ignored. The goal of our paper is to fill this gap.

We contribute to the econometric and empirical literature on Bayesian games in two ways. First, we introduce a general, feasible control-function method for estimating discrete Bayesian games with endogenous states. We model endogeneity in covariates through a triangular system that is flexible enough to accommodate correlation through both player-level and game-level unobserved heterogeneity.[2] We construct control function variables as residuals from auxiliary regressions using exogenous instruments. (In our application of entry in 4G-LTE markets, we use the lagged demographics of neighboring markets as instruments.) We propose a two-step nested pseudo-likelihood (2SNPL) estimator, and show it is root-n consistent and asymptotically normal. Our Monte Carlo simulation shows that the estimator works well in finite samples with moderate sizes.

Heckman (1978), Newey (1987) and Rivers and Vuong (1988) propose methods for dealing with endogenous discrete and continuous covariates in single-agent qualitative response models. While there are other solutions for endogeneity in the literature,[3] the control function approach has proliferated due to its simplicity, flexibility and

[2]Grieco (2014) studied discrete games with exogenous covariates but a flexible information structure that incorporates both game- and individual-level unobserved heterogeneity. In comparison, we use a triangular system to allow for flexible sources of endogeneity in covariates. Marcoux (2020) uses bids for spectrum licenses in Canadian telecommunications industry to recover the unobserved heterogeneity of competitors, and then uses its variation to identify incumbents' responses to the new entrant's decisions. Unlike our work, bids for spectrum licenses are not payoff-relevant states in the game Marcoux studies.

[3]Lewbel (2000), Blundell and Powell (2004), Rothe (2009) and Hoderlein (2014) deal with endogeneity in semiparametric binary choice models; Vytlacil and Yildiz (2007) consider nonparametric identification and estimation of average treatment effects of dummy endogenous variables in weakly separable models; Dong and Lewbel (2015) estimate binary choice models with discrete, continuous, or censored endogenous regressors. D'Haultfœuille and Février (2015) and Torgovitsky (2015) show that non-separable models with continuous outcome and endogenous variables can be identified using discrete instruments.

wide applicability.[4] We contribute to this extensive literature by bringing the control function approach into a game-theoretic setting with incomplete information.[5] We combine control functions with a nested pseudo likelihood method to handle the simultaneity embedded in the Bayesian game and the endogeneity in regressors at the same time.[6]

Our second contribution is empirical. We apply the 2SNPL estimator to analyze a hypothetical 2016 T-Mobile and Sprint merger in the U.S. cellphone service market.[7] In addition to predicting firm entries in post-merger local markets and the population served, we evaluate the impact of adding a fourth national provider, enabled by a government-mandated partial divestiture of assets owned by the merging parties. This is a very meaningful exercise for antitrust and regulatory agencies. For example, while reviewing merger proposals, the Federal Trade Commission (FTC) and the Department of Justice (DOJ) have often mandated that the merging firms divest certain assets and facilities to rivaling firms. The goal of such a policy is to strengthen after-merger competition in local markets and to alleviate the loss of consumer welfare due to increased market power of the merged entity.[8] In the case of the 2020 T-Mobile/Sprint merger case, the DOJ required the merging parties to divest parts of Sprint's prepaid businesses, Sprint's 800 MHz spectrum holding, decommissioned cell sites and retail locations to a potential competitor, DISH Network.

A crucial step in our analysis is to allow for endogeneity in providers' network investment while analyzing their strategic decisions to enter local markets. As noted earlier, two covariates that influence strategic decisions are endogenous (3G deployment in the focal market and 4G-LTE deployment in neighboring markets). Our estimates indicate that unobserved factors in a firm's 4G-LTE deployment decision are

[4]Since its inception by Heckman and Robb (1985), the control function approach has been used in a variety of settings. See, for example, Newey, Powell, and Vella (1999), Chesher (2003), Das, Newey, and Vella (2003), Lee (2007), Florens, Heckman, Meghir, and Vytlacil (2008), Imbens and Newey (2009), Klein and Vella (2010), Petrin and Train (2010), Hahn and Ridder (2011), and Kasy (2011) among others.

[5]For econometric analyses of static Bayesian games, see Aradillas-Lopez (2010), Bajari, Hong, Krainer, and Nekipelov (2010), Florens and Sbaï (2010), Tang (2010), De Paula and Tang (2012), Misra (2013), Wan and Xu (2014), Lewbel and Tang (2015), Aradillas-Lopez and Gandhi (2016), Lin and Xu (2017), Xu (2018), Aguirregabiria and Mira (2019), Lin, Tang, and Yu (2021) and Aradillas-López (2020).

[6]The fixed-point algorithm is typically used to deal with simultaneity of strategic choices in discrete Bayesian games. See Rust (1987), Aguirregabiria and Mira (2002, 2007) and Kasahara and Shimotsu (2012).

[7]T-Mobile and Sprint proposed a merger deal in 2019 and were approved to merge in 2020 after lengthy legal battles surrounding antitrust concerns. In our simulations, we create a hypothetical merger between these two firms by moving the 2020 merger to the end of 2015.

[8]For example, in 2015, the FTC required Albertsons and Safeway to sell 168 stores in 130 local markets as a condition for approving their $9.2 billion merger case.

negatively correlated with its focal market's 3G deployment and positively correlated with its 4G-LTE deployment in neighboring markets. Both correlations are statistically significant, providing evidence for the endogeneity of these two covariates. These covariates are directly impacted by the merger (the new entity owns a union of network facilities of the merging parties). Thus, any sound analysis of the merger's impact needs to start with a valid, endogeneity-proof inference of covariate effects.

Using our endogeneity-proof estimates, we find that the hypothetical T-Mobile and Sprint merger would substantially reduce the overall 4G-LTE deployment across local markets, despite the merged firm became a strong competitor and owned better assets after taking over Sprint's cellular networks. This finding counters a typical pro-merger argument that cost synergies lead to wider cellular coverage and benefit consumers.[9] Moreover, our simulations show that the addition of a fourth national firm, mirroring the DOJ's DISH Network merger remedy through divestiture, would not completely offset the merger's negative impact on the population served. Lastly, we compare the estimation and simulation results with and without taking into account the endogeneity in network investment. This comparison shows that, taking network investment as exogenous, one would over-estimate the total number of entry occurrences but under-estimate the percentage of population under-served under both the merger and the remedy. Ignoring such endogeneity in estimation would, therefore, skew the policy implications for antitrust agencies toward the merger and its proposed remedy.

As our work incorporates endogenous assets in oligopolistic firms' strategic choices, we build on the recent empirical literature in industrial organization that evaluates how merger affects product offerings (Fan, 2013, Wollmann, 2018, Fan and Yang, 2020), quality of service (Elliott, Houngbonon, Ivaldi, and Scott, 2021) and entry (Berry and Waldfogel, 2001, Sweeting, 2010, Li, Mazur, Park, Roberts, Sweeting, and Zhang, 2019, Ciliberto, Murry, and Tamer, 2021, Fan and Yang, 2021). Mergers, in the first place, are consolidations of assets and resources, including production facilities, retail outlets, investments, patents and more. Divestitures are the regulators' responses aimed at counteracting the increased concentration in post-merger assets distribution. Empirical work evaluating the role of divestiture practices in merger cases is scarce, due partly to the lack of data and partly to the lack of a tractable framework to account for the

[9]Our results are consistent with findings in Genakos, Valletti, and Verboven (2018), which used mobile operators' prices and accounting information across 33 OECD countries over a decade to show that both prices and investment per operator increased after a merger and that total industry investment did not change significantly.

endogeneity of assets and divestiture.[10] To the best of our knowledge, our work is the first to evaluate the role of assets and, more importantly, the role of divestitures in firms' strategic choices using a game-theoretic approach. More broadly, our empirical method provides a very feasible solution to covariate endogeneity in discrete Bayesian games.

The paper is organized as follows. Section 2 introduces the discrete Bayesian games with endogeneity and characterizes the Bayesian Nash equilibrium. Section 3 describes the two-step nested pseudo likelihood estimator (2SNPL) and derives its asymptotic properties. Section 4 illustrates the finite-sample performance of the 2SNPL method using two Monte Carlo experiments. Section 5 studies the 4G-LTE entry game of AT&T, Verizon, T-Mobile and Sprint, comparing model estimates and policy implications with and without accounting for endogenous covariates. Section 6 concludes. All proofs, technical details, and robustness checks are rendered in the appendices.

2 Discrete Bayesian Games with Endogeneity

Consider a Bayesian game of simultaneous discrete choices among K players, indexed by $k \in \mathcal{K} \equiv \{1, 2, ..., K\}$. Each player k is characterized by a $d_x \times 1$ vector of covariates X_k. For each player k, let $X_k \equiv (X_{k1}, X_{k2})$ be partitioned into $d_1 \times 1$ vector of exogenous covariates X_{k1} and $d_2 \times 1$ vector of endogenous covariates X_{k2}. Z_k is $d_z \times 1$ vector of instruments. Each player k observes a private shock $u_k \in \mathbb{R}$, and makes a simultaneous entry decision $Y_k \in \{0, 1\}$ based on the public information $\mathbb{I} \equiv \{X_{k1}, X_{k2}\}_{k \leq K}$ and the private shock u_k.

A player k's ex post payoff for $Y_k = 1$ is

$$X'_{k1}\beta_k + X'_{k2}\gamma_k + \alpha_k \sum_{j \neq k} Y_j + u_k, \tag{1}$$

and that for $Y_k = 0$ is normalized to be zero. The instruments Z_k do not enter the ex post payoffs, but contribute to the endogenous variables as follows:

$$X_{k2} = \Pi'_k (X'_{k1}, Z'_k)' + V_k, \tag{2}$$

where Π'_k is a $d_2 \times (d_z + d_1)$ matrix of constant coefficients. Instrument validity requires the coefficients for Z_k in the matrix Π_k to be non-zero. The regressor X_{k2} is endogenous

[10] Two recent academic papers provide descriptive evidence on the effects of divestitures: Tenn and Yun (2011) compare pre- and post-divestiture performances of divested brands from the 2008 Johnson & Johnson's acquisition of Pfizer's consumer health division; Soetevent, Haan, and Heijnen (2014) evaluate the effects of the Dutch government's divestiture requirement when allocating rights to operate highway gasoline stations on prices of divested gasoline stations.

whenever the error terms $V_k \in \mathbb{R}^{d_2}$ and $u_k \in \mathbb{R}$ are correlated.[11]

In any pure-strategy Bayesian Nash equilibrium (psBNE), each player k follows a decision rule $Y_k = 1\{Y_k^* > 0\}$:

$$Y_k^* \equiv X_{k1}'\beta_k + X_{k2}'\gamma_k + \alpha_k \sum_{j \neq k} \mathbb{E}_k(Y_j|\mathbb{I}, u_k) + u_k, \tag{3}$$

where $\mathbb{E}_k(Y_j|\mathbb{I}, u_k)$ is player k's belief about others' decisions, which is consistent with the common prior of $\{u_j\}_{j\leq K}$ and others' strategies in equilibrium. (We can generalize by letting the strategic interaction term be a weighted sum of other players' choice probabilities — that is, by allowing the weights $\alpha_{k,j}$ to differ across k as well as j.)

Our method for dealing with endogenous covariates in this model applies under intuitive conditions on the unobserved errors, which are formalized as follows. For each k, let η_k denote the error term in the linear projection of u_k on $\{V_j\}_{j\leq K}$. That is,

$$u_k = \sum_{j\leq K} V_j'\lambda_{k,j} + \eta_k, \tag{4}$$

where $\lambda_{k,j}$'s are coefficients in the linear projection. Basically, u_k absorbs all factors that affect firms' *ex post* payoffs but are not reported in the data. While u_k is the structural private shock for player k in the game, η_k is the "residual private information" after projecting u_k onto the vector of V_j's, which can be recovered from commonly observed variables using Equation (2). In this sense, V's of all firms are effectively public information, as in principle they can be backed out by the econometrician from the auxiliary equation that explains the source of endogeneity.

Assumption 1. *(i) $\{u_k, V_k\}_{k\leq K}$ are independent of $X_1 = \{X_{k1}\}_{k1\leq K}$ and $Z = \{Z_k\}_{k\leq K}$ with zero means. (ii) $\{\eta_k\}_{k\leq K}$ are independent of $\{V_k\}_{k\leq K}$; and η_k's are independent across the players $k = 1, 2, ..., K$.*

Part (i) of the assumption states the instrument exogeneity; part (ii) posits that the residual private information, from the projection of u_k on the public control variables, are independent between the individual players.

It is worth emphasizing that this assumption is flexible enough to accommodate different forms of endogeneity in X_{k2}, including those due to player- or game-level unobserved heterogeneity. For example, it allows for existence of information

[11] The linearity in the auxiliary equation (2) is not essential for the method we propose. The control function method applies even when the linear index of (X_{k1}, Z_k) on the right-hand side of equation (2) is replaced by a nonlinear function of (X_{k1}, Z_k), provided the control function variables V_k satisfy Assumption 1. In fact, we use a quadratic functional form with pairwise interactions of all covariates in our application for better fit in the first stage.

commonly known to the players but not observed by researchers in the sample. Consider a data-generating process whereby V_j is arbitrarily correlated across $j = 1, 2, ..., K$, possibly through some game-level unobserved heterogeneity. Suppose that, for each player k, u_k is a linear combination of $\{V_j\}_{j\leq K}$ and an idiosyncratic noise ϵ_k, with ϵ_k being independent across $k = 1, 2, ..., K$, and jointly independent of $\{X_{j1}, Z_j, V_j\}_{j\leq K}$. Then, for all k, the error term from a linear projection of u_k on $\{V_j\}_{j\leq K}$, denoted as η_k, is identical to ϵ_k. Therefore, the conditions in Assumption 1 are satisfied.

Assumption 1 also accommodates situations in which endogeneity arises because of unreported individual heterogeneity. For example, suppose that there is no unobserved heterogeneity on the game level, and $\{u_k, V_k\}_{k\leq K}$ are independent across all players. For each player k, the vector of individual noises (u_k, V_k') is multivariate normal with non-zero correlation between u_k and the components in V_k due to some unobserved characteristics of player k. Assumption 1 follows from an implication of the multivariate normality. The zero mean restriction in (i) is just a location normalization.

Using Assumption 1, we write the decision rule in (3) as

$$Y_k = 1\{Y_k^* > 0\} = 1\Big\{X_{k1}'\beta_k + X_{k2}'\gamma_k + \alpha_k \sum_{j\neq k} \mathbb{E}_k(Y_j|\mathbb{I}) + \sum_{j\leq K} V_j'\lambda_{k,j} + \eta_k > 0\Big\}. \quad (5)$$

Note that the two conditions in Assumption 1 imply that η_k is independent of $\{X_j, V_j\}_{j\leq K}$ and, consequently, from $\mathbb{I}$. Besides, the independence of η_k across $k = 1, 2, ..., K$ implies that the equilibrium belief $E_k(Y_j|\mathbb{I}, u_k)$ does not depend on the residual private information $\{\eta_j\}_{j\neq k}$.[12]

Let F_k and f_k denote the marginal distribution and the density function of η_k, respectively. Thus, we characterize a psBNE through a vector of conditional choice probabilities (CCPs) $P : \mathbb{I} \mapsto [0, 1]^K$ that solves a fixed-point equation:

$$P = \Gamma(\theta, P), \quad (6)$$

where $\Gamma \equiv (\Gamma_1, ..., \Gamma_K)'$ with

$$\Gamma_k(\theta_k, P) \equiv F_k\Big(X_{k1}'\beta_k + X_{k2}'\gamma_k + \alpha_k \sum_{j\neq k} P_j + \sum_{j\leq K} V_j'\lambda_{k,j}\Big), \quad (7)$$

and $\theta \equiv \{\theta_k\}_{k\leq K}$ with $\theta_k \equiv (\gamma_k', \beta_k', \alpha_k', \lambda_k')$.

Assumption 2. *The sample is generated from a single equilibrium.*

Assumption 2 is common in the estimation of semiparametric Bayesian games

[12] To see this, note that conditioning on $\mathbb{I}$ and u_k is equivalent to conditioning on $\{X_k, V_k\}_{k\leq K}$ and η_k, and the claim follows from the independence of η_k across k.

that exploits the conditional choice probabilities in equilibrium. Examples include Aradillas-Lopez (2010), Bajari, Hong, Krainer, and Nekipelov (2010) and Tang (2010).

Let P^* denote the profile of conditional choice probabilities in a Bayesian Nash equilibrium. Identification using Equations (6) and (7) requires the usual rank conditions. That is, the support of the vector $(X'_{k1}, X'_{k2}, \sum_{j \neq k} P^*_j, V'_1, \cdots, V'_K)$ is not contained in a linear subspace. Note that this rules out the cases in which the coefficients for Z_k in the matrix Π_k are all zeros. This requires a necessary order condition that there are more instruments in Z_k than endogenous variables in X_{k2}.

By Assumption 1-(i), $\{V_k\}_{k \in \mathcal{K}}$ can be recovered directly as the residuals in the regression of X_{k2} on (X_{k1}, Z_k) in Equation (2), and therefore treated as known covariates for subsequent identification. Aradillas-Lopez (2010) and Bajari, Hong, Krainer, and Nekipelov (2010) provide two distinct sets of conditions under which the players' ex post utility functions are identified.

We conclude this section by noting that identification of this model using our control function (CF) approach does not require parametric assumptions on the distribution of η_k. For estimation, one can apply the semiparametric estimators in Aradillas-Lopez (2010) to the reduced form in equation (5) by plugging in estimates of V_k. Hence, the CF method we propose is more robust than parametric alternatives (which would require fully parametrizing the conditional distribution of u_k given all endogenous covariates X_{k2}). The stochastic restrictions on the unobserved errors we maintain in Assumption 1 are nonparametric, and have direct structural interpretation.

3 Estimation

Consider a sample of n independent games $i = 1, 2, ..., n$, each involving K players making simultaneous binary decisions. Throughout this section, we use lower-case letters to denote realization of random vectors in the sample. In each game i and for each player k, the sample reports a binary choice $y_{k,i}$, endogenous variables $x_{k2,i}$, and exogenous covariates and instruments $(x_{k1,i}, z_{k,i})$. Let $\mathbb{I}_i = \{x_{k,i}, z_{k,i}\}_{k \leq K}$ denote the information set that is common knowledge shared by all players in a game.

Let Θ and $\mathcal{P} \subseteq [0, 1]^{K \times |\mathcal{X}| \times |\mathcal{Z}|}$ denote the parameter spaces for θ and P, respectively, with $\mathcal{X}, \mathcal{Z}$ being marginal support of X_k, Z_k. Let $\theta_0 \in int(\Theta)$ denote the true value of θ in the data-generating process (DGP), and let $P^0 \equiv \{Pr\{Y = y | \mathbb{I} = (x, z)\} : (y, x, z) \in \{0, 1\}^K \times \mathcal{X}^K \times \mathcal{Z}^K\}$ denote the actual equilibrium choice probabilities given θ_0 in the DGP.

Assumption 3. *(i) For any $\theta \neq \theta_0$ and $P(\theta)$ that solves $P = \Gamma(\theta, P)$, $P(\theta) \neq P(\theta_0) \equiv P^0$; (ii) common knowledge variables X_i and Z_i have finite supports, denoted as $\mathcal{X}$ and $\mathcal{Z}$; (iii) $(Y_i, X_i, Z_i)_{i=1}^n$ are independent across games, and $Pr\{\mathbb{I}_i = (x,z)\} > 0$ for all $(x,z) \in \mathcal{X}^K \times \mathcal{Z}^K$.*

Assumption 3(i) is a standard identification condition for estimating games in which the equilibrium is characterized by the solution to a fixed-point problem. See, for example, Assumption 5(C) in Aguirregabiria and Mira (2007) and Assumption 1(e) in Kasahara and Shimotsu (2012). Other papers on asymptotic properties of nested pseudo likelihood estimators in discrete games also assume finite support of states — e.g., Assumption 4 in Aguirregabiria and Mira (2007) and §2.1 in Kasahara and Shimotsu (2012).

We propose a two-step nested pseudo likelihood (2SNPL) estimator that builds on a sequential algorithm combining the nested pseudo likelihood estimator in Aguirregabiria and Mira (2007) with the two-stage conditional maximum likelihood in Rivers and Vuong (1988). The *pseudo likelihood* is:

$$L_n(\theta, P; \Pi) = \frac{1}{n} \sum_{i=1}^{n} l_i(\theta, P; \Pi),$$

where $l_i(\theta, P; \Pi) \equiv \sum_{k=1}^{K} log f_{k,i}(\theta, P; \Pi)$, with $\Pi \equiv \{\Pi_k\}_{k \leq K}$ and $f_{k,i}(\theta, P; \Pi)$ defined as

$$Pr\{x'_{k1,i}\beta_k + x'_{k2,i}\gamma_k + \alpha_k \sum_{j \neq k} P_j + \sum_{j \leq K} (x_{j2,i} - \Pi'_j(x'_{j1,i}, z'_{j,i}))' \lambda_{k,j} + \eta_{k,i} \geq (<) 0\}$$

if $y_{k,i} = 1$ ($y_{k,i} = 0$). Note that in the definition of $f_{k,i}$, the probability measure relates to the marginal distribution of $\eta_{k,i}$, and (x_k, z_k) are fixed realizations.[13]

With a slight abuse of notation, we let $\Gamma(\theta, P; \Pi)$ denote the mapping $\Gamma(\theta, P)$ as defined in Equation (7) when V_j is replaced by its identifiable counterpart $X_{j2} - \Pi'_j(X'_{j1}, Z'_j)'$. This emphasizes how the mapping depends on the first-stage parameter Π.

Our 2SNPL estimator is defined as follows. In the first stage, regress $x_{k2,i}$ on $(x_{k1,i}, z_{k,i})$ to estimate $\widehat{\Pi}_k$ for each $k \leq K$. In the second stage, plug $\widehat{\Pi} \equiv \{\widehat{\Pi}_k\}_{k \leq K}$ into an iterative algorithm in Aguirregabiria and Mira (2007) to construct a *2SNPL sequence* of estimators as follows:

Step 1. Pick an initial guess $\widehat{P}_0$ for P^0. For example, one can obtain such an initial guess from a reduced-form Probit regression.

[13] The term "pseudo likelihood" is used because the argument P in L_n is a generic profile of choice probabilities, rather than the equilibrium choice probabilities P^0.

Step 2. For each $s \geq 1$, calculate an s-stage estimator for θ as

$$\widehat{\theta}_s = \arg\max_{\theta\in\Theta} L_n(\theta, \widehat{P}_{s-1}; \widehat{\Pi}), \tag{8}$$

and update the choice probabilities recursively as

$$\widehat{P}_s = \Gamma(\widehat{\theta}_s, \widehat{P}_{s-1}; \widehat{\Pi}). \tag{9}$$

If the initial guess $\widehat{P}_0$ is a consistent estimator for the actual P^0 in the DGP, then all elements in the sequence of estimators are consistent for θ_0. This follows from a similar argument for the consistency of two-step pseudo maximum likelihood estimators in Proposition 1 of Aguirregabiria and Mira (2007).

More importantly, there exists a neighborhood around P^0 such that, starting from any initial guess $\widehat{P}^0$ in that neighborhood, the NPL sequence constructed above converges almost surely to a root-n consistent and asymptotically normal (CAN) estimator, which we refer to as a *2SNPL estimator* and characterize in the next paragraph.

Define a *2SNPL operator* associated with the iterations in (8) and (9):

$$\phi_n(P) \equiv \Gamma(\tilde{\theta}_n(P), P; \widehat{\Pi}), \text{ where } \tilde{\theta}_n(P) \equiv \arg\max_{\theta\in\Theta} L_n(\theta, P; \widehat{\Pi}). \tag{10}$$

The set of *2SNPL fixed points* in a sample is defined as $\Lambda_n \equiv \{(\breve{\theta}, \breve{P}) \in \Theta \times \mathcal{P} : \breve{P} = \phi_n(\breve{P}) \text{ and } \breve{\theta} = \tilde{\theta}_n(\breve{P})\}$. If the maximizer $\tilde{\theta}_n(P)$ is unique for any P and $\widehat{\Pi}$ from a given sample, then the mapping $\tilde{\theta}_n$ is continuous by the theorem of maximum. Thus, the 2SNPL operator $\phi_n(\cdot)$ is continuous in the compact and convex set $[0,1]^{K\cdot|\mathcal{X}|\cdot|\mathcal{Z}|} \equiv \mathcal{P}$. It follows from Brouwer's fixed-point theorem that Λ_n is non-empty. We define a *2SNPL estimator* $(\widehat{\theta}_{2SNPL}, \widehat{P}_{2SNPL})$ as the element in Λ_n that leads to the highest value of pseudo likelihood.

3.1 Asymptotic Properties of the 2SNPL Estimator

Let Π_0 denote the true value of Π in the DGP. For simplicity, we also use Π, Π_0 to denote their own vectorization, in which case Π, Π_0 are $K \times d_2 \times (d_1 + d_z)$ vectors. Define the population counterparts of $L_n, \tilde{\theta}_n, \phi_n$ by

$$L_0(\theta, P) \equiv \mathbb{E}\Big[l_i(\theta, P; \Pi_0)\Big];$$

$$\tilde{\theta}_0(P) \equiv \arg\max_{\theta\in\Theta} L_0(\theta, P)\ ;\ \phi_0(P) \equiv \Gamma(\tilde{\theta}_0(P), P; \Pi_0).$$

The set of 2SNPL fixed points in the population is $\Lambda_0 \equiv \{(\theta, P) \in \Theta \times \mathcal{P} : \theta = \tilde{\theta}_0(P) \text{ and } P = \phi_0(P)\}$. Let $s_{\theta,i} \equiv \nabla_\theta l_i(\theta_0, P^0; \Pi_0)$, and define

$$\begin{aligned} \Omega_{\theta\theta} &\equiv -E\left[\nabla^2_{\theta\theta} l_i(\theta_0, P^0; \Pi_0)\right] = E\left(s_{\theta,i} s'_{\theta,i}\right); \\ \Omega_{\theta P} &\equiv -E\left[\nabla^2_{\theta P} l_i(\theta_0, P^0; \Pi_0)\right] = E\left(s_{\theta,i} s'_{P,i}\right) \text{ where } s_{P,i} \equiv \nabla_P l_i(\theta_0, P^0; \Pi_0); \\ \Omega_{\theta\Pi} &\equiv -E\left[\nabla^2_{\theta\Pi} l_i(\theta_0, P^0; \Pi_0)\right] = E\left(s_{\theta,i} s'_{\Pi,i}\right) \text{ where } s_{\Pi,i} \equiv \nabla_\Pi l_i(\theta_0, P^0; \Pi_0). \end{aligned}$$

The equalities following the definition above are due to the information matrix equality with regard to the vector of scores. We denote the Jacobian matrices evaluated at the true value $(\theta_0, P^0; \Pi_0)$ as $\Gamma^0_P \equiv \nabla_{P'} \Gamma(\theta_0, P^0; \Pi_0)$, $\Gamma^0_\theta \equiv \nabla_{\theta'} \Gamma(\theta_0, P^0; \Pi_0)$, and $\Gamma^0_\Pi \equiv \nabla_{\Pi'} \Gamma(\theta_0, P^0; \Pi_0)$. Define $M \equiv \Omega_{\theta\theta} + \Omega_{\theta P}(I - \Gamma^0_P)^{-1}\Gamma^0_\theta$. We establish the asymptotic property of $\widehat{\theta}_{2SNPL}$ under the following regularity conditions.

Assumption 4. *(i) Θ is a compact convex subset of a Euclidean space, and $\mathcal{P}$ is a compact convex subset of $(0,1)^{n \cdot |\mathcal{X}| \cdot |\mathcal{Z}|}$; (ii) $E\left[\sup_{\theta,P} |l_i(\theta, P; \Pi_0)|\right] < \infty$. (iii) (θ_0, P^0) is an isolated population NPL fixed point (i.e., it is unique, or else there is an open ball around it that does not contain any other element of Λ_0); (iv) there exists a closed neighborhood of P^0, denoted by $\mathcal{N}(P^0)$, such that, for all P in $\mathcal{N}(P^0)$, $L_0(\theta, P; \Pi_0)$ is globally concave, and its second derivative with respect to θ is a nonsingular matrix; (v) the operator $\phi_0(P) - P$ has a nonsingular Jacobian matrix at P^0; (vi) M is nonsingular.*

Recall that $\widehat{\Pi}$ consists of 1st-stage ordinary least squares (OLS) estimates and, therefore, admits a linear, first-order asymptotic representation as

$$\sqrt{n}(\widehat{\Pi} - \Pi_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} r_i(\Pi_0) + o_p(1),$$

where $r_i(\Pi_0) \equiv r_{0,i}$ is the influence function characterizing the limit distribution of the OLS estimator.

Theorem 1. *Under Assumptions 1 and 3 to 5, $\hat{\theta}_{2SNPL}$ is a consistent estimator, and*

$$\sqrt{n}(\widehat{\theta}_{2SNPL} - \theta_0) \xrightarrow{d} N\Big(0, M^{-1} E(\tilde{s}_i \tilde{s}'_i)\left(M^{-1}\right)'\Big),$$

where

$$\tilde{s}_i \equiv s_{\theta,i} - [\Omega_{\theta P}(I - \Gamma^0_P)^{-1}\Gamma^0_\Pi + \Omega_{\theta\Pi}] r_{0,i}.$$

Proof. See Appendix A. □

The proof of the theorem amounts to writing down the first-order conditions and the equilibrium constraints that define the 2SNPL estimator, and then using a first-order expansion to account for the impact of the first-stage estimator $\widehat{\Pi}$, as well as the concurrent iteration over conditional choice probabilities.

The model admits a unique psBNE with primitive condition.

Assumption 5. *For each k, the magnitude of α_k is bounded above: $|\alpha_k| < \frac{1}{(K-1)|\sup_t f_k(t)|}$.*

This assumption restricts the strength of interaction between players so that Γ satisfies the contraction mapping property.[14]

Lemma 1. *Under Assumptions 1 and 5, there exists a unique psBNE.*

Proof. See Appendix A. □

Similar to Kasahara and Shimotsu (2012), we can establish the following convergence property of the 2SNPL sequence.

Theorem 2. *Suppose that Assumptions 1 and 3 to 5 hold and $\Omega_{\theta\theta}$ is nonsingular. There exists a neighborhood $\mathcal{N}$ around P^0 such that, starting from any initial value $\widehat{P}_0 \in \mathcal{N}$, $lim_{s\to\infty}\widehat{P}_s = \widehat{P}_{2SNPL}$ almost surely.*

The contraction mapping property in Lemma 1 implies that $\rho(\Gamma_P^0) < 1$, where $\rho(\cdot)$ is the spectral radius function. The key condition for convergence in Proposition 1 of Kasahara and Shimotsu (2012) holds.[15] With uniform convergence of $L_n(\cdot;\widehat{\Pi})$ to $L_0(\cdot)$ established in the proof of consistency in Theorem 1 (see Appendix A), the proof of Theorem 2 follows from the same steps in Kasahara and Shimotsu (2012) and is, therefore, omitted for brevity.

4 Monte Carlo Evidence

In this section, we illustrate the finite-sample performance of our 2SNPL estimator by several Monte Carlo experiments. We consider four players in the game, each associated with X_1 and X_2, which are drawn from the bivariate normal distribution with mean zero, unit variance, and covariances 0.5. A pair of independent standard normal variates (v, η) were drawn. We consider two cases: homogeneous competitive effects and heterogeneous competitive effects.

[14] Assumption 5 is similar to the Moderate Social Influence (MSI) condition in the interaction game literature (see Glaeser and Scheinkman, 2003, Horst and Scheinkman, 2006). It is used in the discrete game literature (Brock and Durlauf, 2001, Lee, Li, and Lin, 2014, Lin and Xu, 2017, Xu, 2018, Jackson, Lin, and Yu, 2020, Lin, Tang, and Yu, 2021) for the uniqueness of Bayesian Nash equilibrium.

[15] See Section 2.3 of Kasahara and Shimotsu (2012) for more discussion.

4.1 Homogeneous competitive effects

Consider a game with four players indexed by $k = 1, 2, 3, 4$. We generate the error term as $u_k = \lambda v_k + \eta_k$ and the endogenous variable as

$$X_{k2} = \pi_0 + \pi_1 X_{k1} + \pi_2 Z_k + v_k.$$

Let (X_{k1}, Z_k)'s be drawn from bivariate normal distribution with mean $(0, 0)$ and variance-covariance matrix $\begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$. We set the true parameter $(\pi_0, \pi_1, \pi_2, \lambda) = (1, 1, 1, 1)$. The conditional choice probabilities $P^0 = (P_1^*, P_2^*, P_3^*, P_4^*)$ in the BNE are solved by

$$P_k^* = \Phi\Big(\beta_0 + \beta_1 X_{k1} + \gamma X_{k2} + \alpha \sum_{j \neq k} P_j^* + \lambda v_k\Big), k = 1, 2, 3, 4.$$

The decisions are then generated by

$$Y_k = 1\Big\{\beta_0 + \beta_1 X_{k1} + \gamma X_{k2} + \alpha \sum_{j \neq k} P_j^* + u_k > 0\Big\}, k = 1, 2, 3, 4.$$

We set the true parameter $(\beta_0, \beta_1, \gamma, \alpha) = (1, 1, 1, -0.5)$. Each simulation was based on a random sample of (200,400,800) observations and was replicated 1000 times. We report the average biases and the mean squared errors for true parameter $(\beta_0, \beta_1, \gamma, \lambda, \alpha) = (1, 1, 1, 1, -0.5)$ in Table 1.

Table 1: Homogeneous Competitive Effects

	Average Bias				
n	β		γ	λ	α
200	0.042	0.027	0.016	-0.021	0.019
400	0.010	0.009	0.015	-0.006	0.012
800	0.008	0.010	0.003	-0.002	0.007
	Mean Squared Error				
n	β		γ	λ	α
200	0.102	0.040	0.017	0.021	0.024
400	0.048	0.018	0.009	0.011	0.011
800	0.023	0.010	0.004	0.005	0.006

Table 2: Heterogeneous Competitive Effects

	Average Bias					
n	β		γ	λ	α	
200	0.021	0.024	0.028	0.028	-0.010	-0.029
400	0.013	0.016	0.012	0.015	-0.005	-0.015
800	0.000	0.008	0.010	0.006	-0.001	-0.008
	Mean Squared Error					
n	β		γ	λ	α	
200	0.090	0.039	0.017	0.023	0.025	0.024
400	0.041	0.018	0.009	0.011	0.011	0.010
800	0.022	0.009	0.004	0.006	0.006	0.005

4.2 Heterogeneous competitive effects

In this section, we consider Monte Carlo designs in which the competition effects differ across "strong" and "weak" players. All other settings are the same as in the homogeneous case, except that the conditional choice probabilities $P^0 = (P_1^*, P_2^*, P_3^*, P_4^*)$ in BNE are solved by

$$P_k^* = \Phi\Big(\beta_0 + \beta_1 X_{k1} + \gamma X_{k2} + \alpha_k \sum_{j\neq k} P_j^* + \lambda v_k\Big), k = 1, 2, 3, 4,$$

where $\alpha_1, \alpha_2 = \alpha_S$ and $\alpha_3, \alpha_4 = \alpha_W$.[16]

The decisions are then generated by

$$Y_k = 1\Big\{\beta_0 + \beta_1 X_{k1} + \gamma X_{k2} + \alpha_k \sum_{j\neq k} P_j^* + u_k > 0\Big\}, k = 1, 2, 3, 4,$$

where we have the true parameter $(\alpha_S, \alpha_W) = (-0.5, -1)$. Each simulation is based on a random sample of (200,400,800) observations and is replicated 1000 times. We report the average biases and the mean squared errors for $(\beta_0, \beta_1, \gamma, \lambda, \alpha_S, \alpha_W)$ with true values $(1, 1, 1, 1, -0.5, -1)$ in Table 2.

Both Tables 1 and 2 show that our estimator converges to the true parameter values at the parametric root-n rate. In both cases, the variances of the estimators seem to be the dominating component in the mean-squared error (relative to bias).

[16] In this example, we take the first two players as strong and the rest two players as weak. The labels for heterogeneous competition effects indicates what type of these players are, not the type of competitors they will face after entry.

5 Empirical Study: An Entry Game of Cellphone Service Providers

In this section, we illustrate how our method, which takes account of endogenous covariates, provides new insights in policy analyses in a setting where oligopolistic firms compete through strategic 4G-LTE deployment decisions in local markets. The firms are the four national cellphone service providers in the U.S.: Verizon Wireless, AT&T Mobility, T-Mobile US and Sprint Corporation (collectively referred to as the "Big Four").[17] The time period we look into is from 2015 to 2018, a few years before the proposal of a T-Mobile and Sprint merger in 2019, which eventually went through in early 2020 after lengthy legal battles over antitrust concerns.

In this industry, firms make capital investments in cellular networks and transmission facilities before providing services to consumers. Such investments have typically been made in accordance with the dominant technology of the time. For example, throughout most of the 2000s, the third generation of cellphone technology (3G) was the predominant technology, utilizing the 1850 - 1990 MHz spectrum range. Starting from roughly 2010, it was time for the next generation of technology, 4G-LTE.[18] A firm with 3G deployment in a local market can repurpose spectrum used by 3G to support 4G-LTE and can utilize existing facilities, such as cell towers, with upgraded equipment. Such investment also involves heavy spatial consideration. For example, extending coverage from central Phoenix to nearby cities and towns would be easier than providing *de novo* services to these markets. We measure a potential entrant's network investment for a local market by the firm's 3G deployment in the focal market and 4G-LTE deployment in nearby markets. These two sets of network investments are the firm-specific, endogenous covariates we focus on in our empirical framework. They are important determinants of a firm's decision to provide a new generation of technology in a local market, driven by similar unobserved heterogeneity that underlies a firm's entry decision.

[17] We will refer to them as Verizon, AT&T, T-Mobile and Sprint henceforth. These cellphone service providers are also known as mobile network operators, wireless service providers, wireless carriers, cellular companies, mobile network carriers, etc. In this paper, we refer to them as firms, providers, and carriers interchangeably.

[18] 4G-LTE stands for the fourth generation, Long Term Evolution. LTE is the technology to deliver 4G standards, defined as having peak upload and download speeds of at least 100 mbps (mega bits per second). 4G-LTE is still not fully 4G, but is considered the closest to 4G standards by international telecommunications communities.

In the following subsections, we describe the background of the U.S. cellphone service industry, the policy relevance of our empirical application, the data we construct, and the empirical specification we use. In particular, we evaluate a counterfactual experiment in which T-Mobile and Sprint merged in 2016, which would have led to different 4G-LTE deployment paths in markets that these firms had not yet entered in 2016. We discuss the discrepancies in policy implications and recommendations with and without accounting for the endogeneity in network investment.

5.1 The U.S. Cellphone Service Industry at a Glance

Up until April 2020, Verizon, AT&T, T-Mobile and Sprint were the four major cellphone service providers in the United States. There were also a few regional providers, such as U.S. Cellular and C Spire Wireless, and a fringe of local providers, such as Cricket Wireless and TracFone Wireless, which often offered flexible, more economical prepaid plans. Compared to the Big Four, the other providers' network deployment and market presence were almost negligible.[19]

A consumer (or a household) chooses a plan offered by a provider, considering prices, coverage, speed and customer service. A plan typically ranges from $30 to $100. Among the Big Four, Verizon and AT&T were known for the best coverage, while T-Mobile and Sprint were seen as offering comparable deals with lower prices but less coverage. The Federal Communications Commission (FCC) is the main regulator of this industry, while the Department of Justice (DOJ) and the Federal Trade Commission (FTC) share the responsibility for evaluating anti-competitive conduct in this industry.

5.2 Cellular Network Investment

A cellular network is composed of cellphones, base transceiver stations ("cell sites"), mobile telephone switching offices, and the public switching telephone network.[20] When joined together, cellular networks provide radio coverage over a wide geographic area, enabling cellphones to communicate with each other. Globally, major telecommunications providers have deployed cellular networks over most of the inhabited land area on Earth.

[19]The Big Four and US Cellular are the only Mobile Network Operators (MNOs) in the continental U.S. — that is, providers that own and control the spectrum licenses and network infrastructure necessary to provide services to subscribers. All other cellphone providers in the U.S. are Mobile Virtual Network Operators (MVNOs), relying on other firms' network infrastructure to provide services.

[20]We explain the components and evolution history of cellular networks in Appendix B.

Building a cellular network takes decades of physical and financial investment from a provider. In the past few decades, mobile wireless technologies have experienced multiple generations of evolution — namely, from 0G to 5G. In the 2000s, 3G technology was implemented, enabling media streaming with high connection speed. From the start of the 2010s, 4G-LTE was rolled out gradually, accounting for more than half of mobile connections, hitting 52% for the first time in 2019.[21] Cellular networks need to be maintained and updated constantly, with a substantial cost for sustaining network operation. The Global System for Mobile Communications (GSM) Association projected in 2020 that global network operators would invest more than $1.1 trillion in their networks in the next five years.

During our study period, from 2015 to 2018, 4G-LTE grew to be the dominant network technology. The Big Four have constructed their main 4G-LTE networks, but even extending services to an unserved local market from this main network involves millions to billions of dollars. A potential entrant for a local market needs to first acquire spectrum licenses, depending on the size of the market served.[22] A provider then needs to build cell sites, purchase radio transmitters and receivers, and acquire access to intermediate links connecting different wired networks ("backhaul"). The firm also must build a distributional network and market its services to retail consumers. To sum up, the biggest hurdle of deploying a new network technology is the substantial costs involved; these costs can become prohibitive in areas with low population density and rugged terrains. Retiring technologies of previous generations can free up spectrum and existing facilities to accommodate the next generation of technology; at the same time, deploying a new technology in a cluster of nearby markets, simultaneously or sequentially, helps a provider to achieve economies of scale. For these reasons, it is essential to incorporate the "network investment" effect in a potential entrant's evaluation of the expected payoff from entering a local market. When we study providers' decisions to enter local markets, not accounting for the network investment factor means ignoring a first-order difference between Verizon, an industry leader, and Cricket Wireless, a fringe player.

[21] Industry experts predict that 4G will peak at just under 60% by 2023 (The GSM Association Intelligence, "The Mobile Economy 2020.")

[22] A spectrum license gives its holder the exclusive option to use a certain range of frequencies in a well-delineated geographic area. A firm can purchase these licenses in the FCC spectrum auctions or acquire them in secondary markets through purchase or renting. Xiao and Yuan (2021) describe the 2008 FCC auction to sell off 700 MHz, used mainly for 4G-LTE deployment.

5.3 T-Mobile and Sprint Merger: Policy Considerations

T-Mobile and Sprint announced a merger deal of $26 billion on April 29, 2019. The proposed merger would reduce the number of national providers from four to three, leading to antitrust concerns by state governments and regulating agencies.[23] The merging parties claimed a substantial saving of $43.6 billion via cost synergies, which would allow the merged firm to become a stronger competitor against Verizon and AT&T. Proponents of this merger argued that it would generate broader coverage, greater capacity, higher service quality and a rapid deployment of a nationwide 5G network (Wallsten, 2019). Opponents argued that the reduction in the number of providers would lead to higher prices, fewer choices, lower quality, and a slow rollout of 5G services.[24]

On July 26, 2019, the DOJ approved the merger after T-Mobile and Sprint reached an agreement to sell Sprint's branded prepaid business,[25] Sprint's entire 800 MHz portfolio, and other assets to the DISH Network ("DISH" henceforth). The DOJ believed that DISH's previous spectrum holdings and the divested assets from the merger would help DISH become the fourth national provider. The DOJ also prescribed detailed operational instructions for DISH to enter as a facilities-based provider instead of just a reseller.[26] The DOJ argued that this remedy would restore the *ex ante* competitive market conditions before the merger. Judge Victor Marrero of the U.S. District Court cited the DOJ's remedy as a key factor in approving the merger, noting that it made Dish "well poised to become a fourth [Mobile Network Operator] in the market, and its extensive preparations and regulatory remedies indicate that it can sufficiently replace Sprint's competitive impact." However, opponents questioned the effectiveness of this remedy, calling it "exceedingly optimistic" (Economides et al., 2019) or stating "the Court may have erred in treating DISH as a merger-induced entrant" (Caradonna, Miller, and Shue, 2021).

[23]Internationally, the telecommunications industry has experienced a wave of consolidation activities recently. Most notably, the European Commission allowed four-to-three mergers in the Netherlands, Austria, Ireland, Germany and Italy, but blocked a similar merger in Denmark (Genakos, Valletti, and Verboven, 2018).

[24]DOJ Complaint, U.S. et al. v. Deutsche Telekom AG, T-Mobile Us, Inc., Softbank Group Corp., and Sprint Corporation, No. 1:19-cv-02232, at 3 (D.D.C. Jul. 26, 2019) Case 1:19-cv-02232, July 26, 2019.

[25]This includes Boost Mobile and Virgin Mobile, representing 9.3 million consumers.

[26]The DOJ imposes on the merging parties an obligation to permit DISH to operate as a reseller on the merged firm's wireless network for the entire seven-year term of the settlement. DISH promised to comply with the network build commitments made to the DOJ by 2023. If DISH's own network does not serve 70% of the country by then, it will face penalties of up to $2.2 billion.

On October 18, 2019, the merger received formal approval from the FCC in a 3-2 commissioner vote, but attorney generals from 14 states soon filed lawsuits to block the merger. After lengthy negotiations with the states and the DOJ, the merger officially closed on April 1, 2020, with the Sprint brand discontinued on August 2, 2020.

Evaluating the overall effects of the merger is beyond the scope of this paper. Instead, we focus on evaluating a key claim of the merger's benefit: it would strengthen competition in rural areas and alleviate the divide in cellular infrastructure across the states (Wallsten, 2019). The pre-merger T-Mobile and Sprint did not have sufficient assets and coverage to compete effectively with the industry leaders, especially in rural areas.[27] The merged firm, aided by "the unique combination of spectrum, sites and equipment of T-Mobile and Sprint",[28] would become a comparable rival to AT&T and Verizon. Opponents of the merger, such as the Rural Broadband Association, argued that T-Mobile had shown little incentive to invest in rural areas, and, therefore, its incentives were unlikely to change following this merger.

We investigate how a hypothetical T-Mobile and Sprint merger in 2016 would have affected the 4G-LTE deployment in open markets by national providers. As discussed above, cellphone coverage in unserved and underserved markets is a major policy consideration evaluating the 2020 merger case. No direct empirical evidence, however, is available to support either side of the argument. We also evaluate the remedy proposed by the DOJ, which divests assets from the merger to support DISH as a national provider. We exploit data and a structural model of discrete Bayesian games to analyze the impact of the hypothetical merger and remedy, taking into account the firms' post-merger network consolidation and strategic responses.

5.4 Data Sources

We use three publicly available data sets to construct our sample. The first data are from the FCC's Mobile Deployment Form 477 from 2015 to 2018, which reports semi-annually each provider's 2G-4G coverage in every U.S. census block.[29] The FCC requires all facilities-based broadband providers to file Form 477, which discloses

[27]The FCC reported that in December 2016, more than 98% of rural Census blocks had at least one LTE provider, but only 57% had at least four providers, compared to 96% of non-rural blocks.

[28]T-Mobile and Sprint, "Description of Transaction, Public Interest Statement, and Related Demonstrations", June 18, 2018, page 16.

[29]The FCC started to report the Mobile Deployment (including both voice and broadband) data from December 2014, but 2015 was the first year that the FCC reported the actual area coverage within a census block by each provider. Much of the information presented on data description is based on the FCC's Public Notice (DA 16-1107), released on September 30, 2016.

where they offer Internet access service at speeds exceeding 200 kbps in at least one direction. In particular, for each mobile network technology deployed in each radio frequency band, facilities-based mobile providers must submit polygons representing their nationwide coverage area of that technology and the advertised data upload and download speeds. Providers' submission of data is mandatory, and they must certify the accuracy of the data submitted.

With providers' submitted data on coverage polygons, the FCC reports the percentage of the area in a census block covered by each technology (including 2G, 3G, 4G-non-LTE,[30] and 4G-LTE) by each provider, using a computationally intensive process.[31] In addition, the FCC reports the percentage of a census block covered by "any" technology. From December 2015 to December 2018, the FCC data provide seven snapshots of each firm's granular-level network deployment information. Each snapshot of data has about 45 million observations at the firm-census block level.

The second data set is the 2016 American Community Survey. We obtain demographic variables such as population size, age, gender and ethnicity profiles, income, and commuting patterns that are potential determinants of a consumer's cellphone use. The third data set is the 2000 Population Census. We use exactly the same variables as the ones we obtain from the 2016 American Community Survey, to be used to construct our instrument variables – the lagged demographics of neighboring markets – for endogenous network investment variables. Using information from neighboring markets as instruments is common in industrial organization literature, (e.g. Nevo, 2001). For both data sets, we obtain demographic variables at the census tract level.

5.5 Variable Definition and Sample Construction

With the raw data, we define open markets for 4G-LTE deployment by the four major national providers and then merge in demographic variables at the census tract level.

[30] 4G-non-LTE refers to technologies that do not reach 4G standards but were marketed as 4G by cellphone providers. 4G-non-LTE will be ultimately replaced by 4G-LTE. Sprint and Clearwire, for example, invested in WiMax rather than LTE and had to rebuild their 4G networks.

[31] The FCC first removes the spectrum and speed information from each shapefile filed by a provider, and then consolidates different polygons for a particular technology for a particular provider into a single, unique polygon. The FCC then determines how much of a census block is covered by this unique polygon. The FCC has not calculated how much the coverage reported for one technology does or does not overlap with coverage of another technology — e.g., 2G and 3G overlap within a census block.

5.5.1 Aggregation to census tracts

We use the December editions of the FCC's Mobile Deployment Form 477 data from 2015 to 2018, which yield a four-year snapshot of mobile network deployment for the universe of U.S. census blocks. A census block is the smallest geographic unit in the U.S. Census, amounting to more than 11 millions observations in the 2010 Census. A census block is typically a very small geographic area; for example, it is often a city block bounded on all sides by streets, and we do not think that deployment decisions are made on such a fine-grained geographic basis. We therefore aggregate these census blocks to a universe of 73,057 census tracts. A census tract is designed to be a relatively homogeneous unit with respect to population characteristics, economic status and living conditions. In general, each census tract encompasses 2,500 to 8,000 people.[32] A census tract usually covers a contiguous area; however, the spatial size of census tracts varies widely depending on the density of the settlement. A rough estimate of the radius of a typical census tract is 6.5 kilometers.[33] Although cell towers have a maximum range of 50 to 70 kilometers, they are typically spaced two to three kilometers apart to adequately handle cellphone traffic.[34] Based on the above comparison, we define census tracts as geographic markets based on which cellphone providers make investment and network deployment decisions.

For every firm in every census tract, we calculate the percentage of census blocks covered within the census tract by a given technology. The FCC-reported census block coverage has a bipolar distribution, with a small peak between 0% and 10% coverage and a major peak at 100%.[35] We think that some census blocks may experience low, spillover coverage from a nearby cell site in another census block, and this is not an actual entry. Therefore, we define a provider's coverage of a census block under a given technology as a dummy variable that equals 0 if the FCC-reported coverage falls below 10%, and 1 otherwise. When we aggregate to census tracts, we use the same reasoning and define the entry dummy for 4G-LTE as 0 if the percentage of census blocks covered by 4G-LTE within a census tract falls below 10%, and 1 otherwise.

[32] Due to their size and internal homogeneity, Seim (2006) uses census tracts as location choices for video retail stores.

[33] The total area of the U.S. is 9.857 million square kilometers, covering 73,057 census tracts. A census tract covers 134.9 square kilometers, on average, with roughly 6.5 kilometers as the radius.

[34] In urban areas, cell towers may be 400 to 800 meters apart to accommodate the dense population.

[35] For example, for AT&T and Verizon, the 4G-LTE coverage of a census block was already 100% at the 10th percentile for most of our data period; for T-Mobile and Sprint, this number was at high 90% at the 10th percentile.

5.5.2 Sample construction

By December 2015, 4G-LTE has been well deployed by the Big Four across the U.S., although the fringe competitors (about 80 of them in total) lagged distantly behind. Verizon had entered 98.8% of the 73,057 census tracts; AT&T followed closely with 98.1%; and T-Mobile and Sprint trailed them, with 95.3% and 92.4%, respectively. The non-big-four firms had much smaller coverage in comparison. Even the largest one, U.S. Cellular, had entered only 13.9% of the 55,644 census tracts for which it could have been considered a potential entrant.[36] Other fringe competitors' entry rates were typically below 5% in any given year, and they often considered only urban markets or urban clusters in rural areas.

For the Big Four, we can safely argue that the national 4G-LTE network was mostly laid out by the end of 2015, and the remaining task was about the leftover, often isolated, open markets. We focus on their decisions to enter these local, isolated markets. For each provider, a census tract is defined as an open market for 4G-LTE deployment (an entry decision) if the deployment dummy was 0 in December 2015. We measure each provider's network investment via any generation of technology by the end of 2015. We then use the 2018 data to measure 4G entry into the open markets, treating the time between 2016 and 2018 as a single period in the cross-sectional data. We decide to focus on the leftover, often isolated, open markets at the end of 4G-LTE deployment and to lump a three-year period into a cross section in order to alleviate the concerns that firms make forward-looking decisions in a dynamic oligopoly game. We think our choice of a static entry game framework captures the first-order strategic considerations in this setting.

To summarize, from 2016 to 2018 the Big Four were the main competitors in the U.S cellphone industry, and they were strategically considering whether to enter the few remaining, scattered, markets left open to 4G-LTE deployment. We define a potential entrant to a market as a Big Four provider who had no 4G-LTE deployment in the market by the end of 2015. A potential entrant is observed as having decided to enter a market if it made 4G-LTE deployment by the end of 2018.[37] Our sample consists of 2,582 Census tracts that have at least two potential entrants by the end of 2015.

[36] A firm usually needs to obtain approval from a state before entry (Fan and Xiao, 2015). If a firm had not operated in a single census tract in a state, we do not consider this firm as a potential entrant to any census tracts of the state. Verizon and AT&T had operated in all states (including the District of Columbia); T-Mobile had entered 50 states and Sprint 49 states.

[37] We keep the other fringe competitors when counting the number of incumbents in a census tract.

5.5.3 Summary statistics: the Big Four's cellphone deployment

In Table 3, we present summary statistics of the Big Four's cellphone technology deployment in their open 4G-LTE markets by the end of 2015. Of the 2,582 census tracts in our sample, Verizon had not entered 645 by the end of 2015 (i.e., no 4G-LTE deployment by the end of 2015); AT&T, 1,132 markets; T-Mobile, 2,185 markets; and Sprint, 2,182 markets. Table 3 shows how the Big Four differed in their technology mix of 2G, 3G, 4G-non-LTE, and 4G-LTE. From 2016 to 2018, Verizon focused almost completely on 4G-LTE; AT&T retired 2G and pushed for 3G, 4G-non-LTE, and 4G-LTE, with 4G-LTE leading the growth; T-Mobile grew all four technologies, again with 4G-LTE making the largest strides; Sprint never deployed 4G-non-LTE and made relatively small steps compared to its rivals. Of the four technologies, 4G-LTE is the one that experienced the most growth across the board from 2015 to 2018. The 4G-LTE growth is also reflected by the percentage of 4G-LTE coverage in other tracts of the same county (referred to as "neighboring tracts" henceforth) and the number of incumbents offering 4G-LTE in the focal markets.

We use two network investment variables to capture the existing facilities owned by a potential entrant (firm k) in the focal market and nearby areas. The first is the firm's 3G deployment in the focal market by the end of 2015 (we call it $X_{k2,1}$). The second is the firm's 4G-LTE deployment in neighboring tracts by the end of 2018 (we call it $X_{k2,2}$). As we discussed in Section 5.2, different generations of cellphone technologies can share basic facilities (e.g., cell towers), and nearby cell sites reduce the cost of extending the network extra miles (e.g., nearby conduits can be extended to bordering neighborhoods). Therefore, $X_{k2,1}$ and $X_{k2,2}$ are shifters for a provider k's entry decision into the focal market.

A potential entrant's network investment can be measured in different dimensions. For robustness, we use a potential entrant k's deployment via *any* previous generation before 4G-LTE in the focal market by the end of 2015 as $X_{k2,1}$, or its 4G-LTE deployment in neighboring tracts by the end of 2015 as $X_{k2,2}$. We discuss the robustness of our results under different measurements of $X_{k2,1}$ and $X_{k2,2}$ in Appendix B. These robustness checks only lead to marginally different point estimates, and conform to the main conclusion in our estimation and counterfactual analyses.

Table 3: Cellphone Service Coverage (2015-2018), by the Big Four

	2015		2018	
Variable	Mean	S.D.	Mean	S.D.
Verizon: potential entrant to 645 tracts				
% blocks with 2G	0.007	0.052	0.019	0.106
% blocks with 3G	0.005	0.043	0.016	0.094
% blocks with 4G-non-LTE	0	0	0	0
% blocks with 4G-LTE	-	-	0.122	0.233
% blocks 4G-LTE, neighbor tracts average	0.447	0.301	0.534	0.365
# incumbents with 4G-LTE	1.297	1.108	1.964	1.460
Entry with 4G-LTE	-	-	0.267	0.443
AT&T: potential entrant to 1,132 tracts				
% blocks with 2G	0.233	0.357	0	0
% blocks with 3G	0.384	0.414	0.468	0.433
% blocks with 4G-non-LTE	0.357	0.409	0.403	0.424
% blocks with 4G-LTE	-	-	0.362	0.396
% blocks 4G-LTE, neighbor tracts average	0.336	0.372	0.545	0.334
# incumbents with 4G-LTE	1.479	0.877	2.511	1.329
Entry with 4G-LTE	-	-	0.542	0.498
T-Mobile: potential entrant to 2,185 tracts				
% blocks with 2G	0.046	0.163	0.133	0.312
% blocks with 3G	0.011	0.083	0.164	0.315
% blocks with 4G-non-LTE	0.003	0.032	0.243	0.366
% blocks with 4G-LTE	-	-	0.496	0.423
% blocks 4G-LTE, neighbor tracts average	0.206	0.326	0.534	0.326
# incumbents with 4G-LTE	1.876	0.821	3.011	1.195
Entry with 4G-LTE	-	-	0.648	0.478
Sprint: potential entrant to 2,182 tracts				
% blocks with 2G	0.154	0.314	0.195	0.345
% blocks with 3G	0.147	0.309	0.173	0.326
% blocks with 4G-non-LTE	0	0	0	0
% blocks with 4G-LTE	-	-	0.208	0.358
% blocks 4G-LTE, neighbor tracts average	0.107	0.231	0.259	0.317
# incumbents with 4G-LTE	2.055	0.787	3.286	1.015
Entry with 4G-LTE	-	-	0.293	0.455

Notes: This table is based on 6,244 tract-firm observations (2,582 Census Tracts, two to four potential entrants in each tract). This table reports the Big Four's coverage of census blocks by each generation of technology, summarized over Census tracts each firm has yet to enter with 4G-LTE by the end of 2015.

5.5.4 Summary statistics: market attributes

In Table 4, we compare the market attributes of the census tracts in the sample for our entry game and those of the remaining parts of the country. The most important determinant of entry is population size. Demand for cellphone services depends on market demographics such as gender, age, ethnicity profiles, education, labor force participation, household income and size. Workers' commuting patterns also contribute to the intensity of cellphone use. Lastly, population density, ruralness and the presence of large areas of water can be considered cost shifters for network deployment.

As shown in Table 4, the 2,582 census tracts, which have at least two Big Four potential entrants, are notably different from the rest of the country in all dimensions. They have much smaller populations and very different demographic compositions. They are more rural, more sparsely-populated, poorer and less educated. They spend more time working from home and less time commuting to work. In short, these markets seem to belong to the bottom side of the "digital divide," which refers to the significant disparity in Internet access across different demographic groups and geographic areas in the country.

5.6 Instrumental Variables

To specify our Equations (1) to (4) in this cellphone 4G-LTE entry game application, we reiterate our notation:

- Y_k: potential entrant k's 4G-LTE entry decision;
- X_{k2}: include potential entrant k's 3G deployment in the focal census tract, $X_{k2,1}$, and its 4G-LTE deployment in neighboring tracts, $X_{k2,2}$;
- X_{k1}: tract attributes from 2016 ACS + the number of 4G-LTE incumbents in the focal census tract by the end of 2015;[38]
- Z_k: instrumental variables for $X_{k2,1}$ and $X_{k2,2}$ (all variables summarized in Table 4);
- u_k: unobserved errors in the *ex post* payoffs (Equation (1));

[38] We treat the number of incumbents as predetermined and uncorrelated with the unobserved u_k in the entry payoff equation. An incumbent's entry decision was made earlier, before the realization of a potential entrant's time-varying private shocks.

Table 4: Census Tract Attributes

		Markets to enter		Other markets	
Variable	Definition	Mean	S.D.	Mean	S.D.
Pop(in 000's)	Population in thousands	2.901	1.758	4.414	2.171
% Female	% female in population	0.495	0.045	0.508	0.050
% Senior	% 65 and older in population	0.285	0.266	0.153	0.092
% White	% White in population	0.862	0.209	0.724	0.253
% Black	% Black in population	0.038	0.107	0.142	0.222
% Native	% Native in population	0.044	0.161	0.008	0.035
% Asian	% Asian in population	0.012	0.037	0.049	0.091
% Hispanic	% Hispanic in population	0.072	0.129	0.163	0.215
% College	% above 25, with college degree	0.193	0.092	0.294	0.190
% Labor force	% above 16, in labor force	0.573	0.109	0.631	0.103
% Work home	% above 16, employed, working at home	0.056	0.047	0.045	0.040
% Long comm.	% above 16, employed, commuting for 40+ minutes	0.167	0.106	0.201	0.129
HH income	Median household income in 2016 $, 000's	46.127	14.499	59.641	29.860
HH size	Household size	4.651	5.688	2.915	1.835
Pop density	Population/land area	0.0002	0.0007	0.002	0.005
% Rural	% population in rural area	0.683	0.404	0.190	0.348
Mostly water	If water area ≥ 90%	0.113	0.317	0.0007	0.027

Notes: This table is based on 73,057 Census tracts, which include 2,582 tracts for the final sample we use for estimation and 70,475 tracts for the rest of the data.

- V_k: unobserved errors in $X_{k2,1}, X_{k2,2}$ (Equation (2)).
- η_k: residual private shocks after we control for V_k (Equation (5))

In the above specification, we focus on 4G-LTE competition. For example, we do not consider a provider that had offered only 3G service as of 2015 as a competitor in the 4G-LTE deployment game. We have two endogenous covariates in a potential entrant's expected payoff function: $X_{k2,1}$ and $X_{k2,2}$. In this subsection, we discuss our choice of Z_k, which serves as instrumental variables for $X_{k2,1}$ and $X_{k2,2}$.

In a firm's decision rule (Equation (5)), the unobserved error η_k is a potential entrant's private information. The potential entrant observes its own η_k, but not the others', before deploying different generations of technology in the focal and nearby markets. A good example of this unobserved term is each firm's cost of deployment, maintenance and operating in the focal market. V_k, however, is public information, as it can be backed out from the auxiliary equation that explains the source of endogeneity. A good example of V_k is a firm's spectrum holdings and lease/roaming agreement in the focal market. Valid instruments for X_{k2} need to be excluded from the entry payoff function, to be orthogonal to η_k and V_k, and to enter the technology deployment equation.

For each focal census tract, we use demographics of its "neighbors" (i.e., other tracts in the same county) in 2000 as instruments for $X_{k2,1}$ and $X_{k2,2}$. The 2000 demographic variables of neighboring tracts determine the 3G and 4G-LTE deployment in these neighboring markets, but they do not enter the 4G-LTE deployment of the focal market directly conditioning on the focal market's observables. Furthermore, it is plausible to assume that these neighboring demographics are orthogonal to the unobserved factors determining deployments in focal and neighboring markets (η_k and V_k), once conditional on market-level observables.

One may worry that a potential entrant makes entry decisions on a much larger scale than a census tract, so these instruments will enter the payoff function of the focal market. However, the remaining tracts to enter in 2016 were typically isolated spots with the surrounding tracts well served before the start of our sampling period, as shown in Table 3's summary statistics on the percentage of census blocks covered with 4G-LTE in the neighboring tracts. Therefore, modeling the Big Four providers' post-2015 4G-LTE deployment decisions on the level of local markets serves as a first-order approximation that captures the firms' main strategic concerns.

We choose neighboring markets' attributes in 2000 (instead of in 2016) as our instruments for several reasons. First, 3G technology was actively deployed between 2000 and 2010; therefore, the 2000 Census's market attributes are more relevant to 3G deployment. Second, the correlation between 2000 Census's market attributes and 2016's market-level unobserved heterogeneity is weakened with time, giving us better justification for the orthogonality of the instruments. For example, a firm's spectrum holdings in 2016 (likely sources of V_k) are less correlated with neighboring markets' attributes in 2000, after multiple generations of cellphone technology evolution and secondary market trading. Lastly, as we can reasonably argue that the detailed market-level attributes we include in X_{k1} capture the spatial correlation across census tracts, we abstract away spatial correlation in the unobservables. That is, conditional on X_{k1}, the unobservables η_k and V_k, which captures firm- and market-specific heterogeneity, is not spatially correlated. To check instrument strength and relevance, we test the joint significance of these instruments in Equation (2). For the equation that explains $X_{k2,1}$, the F-statistic is 1.591 with a p-value less than 0.001; for $X_{k2,2}$, the F-statistic is 3.849 with p-value less than 0.001.

5.7 Estimation Results

Among the Big Four, AT&T and Verizon lead in terms of spectrum holdings, network built and customer base. Our baseline specification categorizes AT&T and Verizon as "strong" competitors and T-Mobile and Sprint as "weak." Analogous to our Monte Carlo exercise, we estimate heterogeneous competitive effects based on whether the potential entrant is strong or weak in terms of deployment, relative to competition. We present results treating all four firms as equal competitors in Appendix C. We adopt a specification in which firms share the same coefficients for all covariates (other than the aforementioned heterogeneous competition effects) in ex post payoffs. Thus, for simplicity, we suppress the generic index k in $X_{k2,1}, X_{k2,2}, V_{k1}, V_{k2}$ when reporting our estimation and simulation results.

Table 5 presents estimation results from two models, with and without accounting for endogeneity in X_2 respectively. In the former case, we include all exogenous and instrumental variables, all their squared terms (excluding dummy variables), and the pairwise interaction terms of all regressors in the first stage estimation. In the latter case, all covariates in X_1 and X_2 are considered exogenous in MLE estimation. Using our estimator in Section 3 to allow for endogenous X_2, we get estimates that mostly conform to our expectations. The "expected competition" effects (α_k) are significantly negative, with a stronger negative effect on weak potential entrants. The incumbent effect is also significantly negative. Population size contributes to 4G-LTE entry positively, but the percentages of seniors and Natives, as well as water coverage, act in the opposite direction. The percentage of labor force participation and the population density, surprisingly, contribute to 4G-LTE entry negatively. We conjecture that the markets that some of the Big Four providers had not entered by 2016 may have inherent differences from other markets, in terms of how population density and labor participation affect profitability.

Allowing for potential endogenous X_2's turns out to have a big impact on the estimates of network investment effects. Both models produce significantly positive estimates for the coefficients of X_2's, but ignoring the endogeneity in X_2 underestimates the effect of $X_{2,1}$ while over-estimating that of $X_{2,2}$. The reason for such discrepancies can be attributed to the roles of structural errors (Vs) in the expected entry payoff. These Vs are firm- or market-level heterogeneity. Different Vs may contribute to 3G/4G-LTE deployment in the focal market and 4G-LTE deployment in the neighboring markets in different directions.

As noted earlier, a good example of V's that can lead to these patterns is each firm's spectrum holdings for different generations of cellphone technology. A firm knows its own spectrum holdings and usually has a rough idea of its rivals' spectrum holdings, because the FCC's spectrum auctions are public information and trading/leasing/roaming agreements are often industry knowledge. The spectrum of a certain frequency often best serves a particular generation of cellphone technology and has different suitability for urban, suburban and rural deployment. For example, 700 MHz is considered the right band for 4G-LTE, while 2.5GHz is right for 5G. A firm may have a rich stock of 3G spectrum but a poor stock of 4G-LTE spectrum, simply due to budget constraints.[39] The negative correlation between a firm's 3G and 4G-LTE spectrum holdings in a focal market (which are captured in V_1 and u, respectively) is consistent with a negative coefficient for V_1 in our estimates, which account for endogenous $X_{2,1}$. In addition, this negative correlation also explains the negative bias in the estimated coefficient for X_{21} when its endogeneity is ignored (i.e., 2.024<3.852).

At the same time, note that if a firm owns a 4G-LTE spectrum license for the focal census tract, this license covers at least the entire county due to its indivisible nature.[40] Hence, there is a positive correlation between the 4G-LTE spectrum holdings in the focal and neighboring markets (captured by u and V_2, respectively). This is consistent with a positive coefficient for V_2 in our estimates accounting for endogenous X_2, and it leads to a positive bias in the estimated coefficient for $X_{2,2}$ when endogeneity is ignored (i.e., 3.042>1.642).

5.8 Counterfactual Results: Evaluating the Merger and the Merger Remedy

In this section, we investigate the impact of a hypothetical merger between T-Mobile and Sprint in 2016. In the first scenario, we use the structural estimates from Table 5 to simulate market outcomes under a baseline scenario with no mergers. In the second scenario for simulation, T-Mobile and Sprint are merged into a "strong" competitor with

[39] For example, T-Mobile did not (and still does not) have enough low-band spectrum (600 MHz), which has wider reach and is better suited for rural deployment; instead, it relies on 1,700MHz and 1900MHz for 4G-LTE deployment, which is better suited for urban and suburban areas.

[40] The FCC's smallest coverage for a spectrum license is the Cellular Market Area, which typically covers three to four counties. Even if firms divide spectrum licenses for resale and lease in secondary markets, they do not break down counties (Kavalar, 2014).

Table 5: Estimation Results of the 4G-LTE Entry Game: Structural Coefficients

	Treating X_2 as Endogenous		Treating X_2 as Exogenous	
	(1)		(2)	
Variable	Estimate	Std. Error	Estimate	Std. Error
Pop (in 000's)	0.078***	0.019	0.081***	0.019
% Female	0.361	0.686	0.417	0.758
% Senior	-1.041***	0.388	-1.110***	0.425
% White	-0.430	0.482	-0.170	0.515
% Black	-0.525	0.511	-0.367	0.535
% Native	-1.555***	0.485	-1.611***	0.518
% Asian	0.894	1.333	1.290	1.316
% Hispanic	-0.529*	0.294	-0.185	0.308
% College	-0.007	0.394	-0.063	0.417
% Labor force	-1.023***	0.286	-0.627**	0.314
% Work home	0.251	0.525	0.768	0.588
% Long comm.	0.036	0.283	-0.508*	0.297
HH income	-0.002	0.003	-0.006**	0.003
HH size	-0.045**	0.023	-0.059**	0.026
Pop density	-0.124*	0.074	-0.229***	0.084
% Rural	0.239**	0.093	0.188**	0.092
Mostly water	-1.731***	0.478	-2.418***	0.545
# Incumbents	-0.122***	0.041	-0.225***	0.042
Intercept	0.991	0.695	0.871	0.757
Network Investment Effects				
$X_{2,1}$	3.852***	0.253	2.024***	0.178
$X_{2,2}$	1.642***	0.143	3.042***	0.096
V_1	-1.982***	0.221	–	–
V_2	2.022***	0.209	–	–
Expected Competition Effects				
Strong potential entrant	-0.955***	0.080	-1.084***	0.085
Weak potential entrant	-0.993***	0.063	-1.133***	0.064

Notes: The results are based on 6,244 observations, corresponding to two to four potential entrants for 2,582 Census tracts. Standard errors are obtained from resampling of markets with replacement 1,000 times. Asterisks indicate the statistical significance at the 1% (***), 5% (**), and 10% (*) levels.

integrated T-Mobile and Sprint network (henceforth referred to as a "New T-Mobile").[41]

[41] After the merger, T-Mobile will bridge the two network cores by routing Sprint traffic to the T-Mobile anchor network. An estimated 11,000 Sprint cell sites will be retained to improve capacity and/or coverage in the new network. We implement the after-merger network integration in data by

In the third scenario, we introduce DISH as a new potential entrant. It is modeled as a "weak competitor" that takes over the decommissioned network originally owned by Sprint. That is, in this scenario, the T-Mobile and Sprint merger is mandated to divest assets to the new competitor DISH, enabling DISH's entry as a facilities-based provider. This scenario corresponds to the DOJ's proposed remedy based on anti-trust concerns. We keep all 2,582 open markets in the baseline simulation, which has a combined total population of 17,209,450.[42]

We use the estimated coefficients in Table 5 to simulate the local market entry decisions of Verizon, AT&T and New T-Mobile (and DISH in the third scenario). For comparison, in each scenario, we simulate two sets of outcomes, one with and one without accounting for endogenous X_2. Panel A in Table 6 presents the simulated market entries across different scenarios; Panel B in Table 6 reports the population still underserved (that is, the population with the number of providers less than or equal to one) by the end of 2018 across these scenarios.

Columns (1) to (3) of Table 6 are simulation results under the three scenarios, using structural estimates that account for endogeneity (column (1) of Table 5). Comparing column (1) to column (2), we can see that the T-Mobile and Sprint merger reduces the number of total entry occurrences from 2,901 to 2,232, a 28.6% reduction rate. This leads to a large increase (23.4%) in the underserved population, especially the rural population (29.4%). There are two explanations for such a reduction: First, there are fewer potential entrants on the markets after the merger. Second, the New T-Mobile resulting from the merger is a strong competitor with integrated deployment from Sprint and T-Mobile and, therefore, is more likely to deter entry by the other competitors. As for each firm's entry occurrences after the merger, New T-Mobile would gain sizable ground after the merger (compared to the pre-merger T-Mobile), while AT&T and Verizon would stay roughly the same. Overall, the reduction of total instances of market entry after the merger is due mainly to the fact there would be fewer potential entrants.

---

taking the union or the maximum of T-Mobile's and Sprint's coverage at the census block level. These two methods yield almost identical results because, in census tracts where both firms serve, they serve mostly at 100% coverage. We report the maximum result in the paper.

[42]If either T-Mobile or Sprint was a 4G-LTE incumbent in a census tract in 2015 and the other was a potential entrant, we assume that after the merger, New T-Mobile will re-evaluate the profitability of the market and decide about entry again. In this case, and in the case that Sprint was a 4G-LTE incumbent in a census tract in 2015, the number of incumbents in 2016 will be reduced by one after the merger.

Table 6: Counterfactual Results under Alternative Models

	Treating X_2 as Endogenous			Treating X_2 as Exogenous		
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Entry outcomes						
# markets with	Baseline	4 to 3	DISH	Baseline	4 to 3	DISH
n entrants = 0	514	661	521	539	582	513
n entrants = 1	1,281	1,618	1,225	1,242	1,669	1,236
n entrants = 2	741	295	772	753	323	770
n entrants = 3	46	8	62	47	8	59
n entrants = 4	–	–	2	1	–	4
Total # entry occurrences	2,901	2,232	2,963	2,893	2,339	2,969
by AT&T	705	713	708	698	747	724
by Verizon	147	154	147	143	183	153
by T-Mobile/New T-Mobile	1,237	1,365	1,296	1,205	1,409	1,262
by Sprint/DISH Network	812	–	812	847	–	830
Panel B: Population (in 1000's) under-served (# incumbents in 2018 ≤ 1)						
Underserved population	Baseline	4 to 3	DISH	Baseline	4 to 3	DISH
Total population	209	258	207	229	244	204
Minority population	125	139	126	136	131	124
Rural population	143	185	155	162	176	151

Notes: The above results are based on 2,582 census tracts, each with one to four potential entrants. In Panel A, the number of entrants does not include the number of incumbents prior to 2016; in Panel B, the number of incumbents in 2018 includes both the incumbents prior to 2016 and the entrants between 2016 and 2018.

Table 7: Merger Evaluation under Alternative Models

	4 to 3 merger effects = (4 to 3 - baseline)/baseline		DISH remedy effects = (DISH - baseline)/baseline	
	(1)	(2)	(3)	(4)
	Endogenous X_2	Exogenous X_2	Endogenous X_2	Exogenous X_2
# unserved markets	28.6%	7.9%	1.4%	-4.8%
# monopoly markets	26.3%	34.4%	-4.3%	-0.5%
Total # entry occurrences	-23.1%	-19.1%	2.1%	2.6%
Total pop. underserved	23.4%	6.5%	-1.0%	-11.0%
Minority pop. underserved	11.2%	-3.7%	0.8%	-8.8%
Rural pop. underserved	29.4%	8.6%	8.4%	-6.8%

Notes: The above results are calculated using Table 6 results.

Now, consider the scenario in which DISH is introduced after the merger as a fourth competitor, enabled by the mandated divestiture from the New T-Mobile. We assume the most optimistic scenario, which is that DISH is able to achieve Sprint's deployment in 2015 after entry. Column (3) of Table 6 reports that the number of entry occurrences is 2,963, substantially higher than in the scenario without the divestiture required by the DOJ. This practically restores the level of market entries before the merger. On most dimensions of entry outcomes and population covered, the DISH remedy seems to be effective. The exception is about the rural population underserved — the rural population underserved would increase by 8.6% from column (1) to column (3), despite that every firm increased entry occurrences or at least remained the same. We checked entry patterns of each firm, and discovered that firms would choose to enter different census tracts after the (remedied) merger, leading to a change in the composition of markets served. In this simulation, we have DISH assuming exactly the same network deployment as Sprint in 2015, and AT&T and Verizon staying the same as before. T-Mobile is the only firm with big changes: New T-Mobile not only has a stronger network, but also becomes a "strong" potential entrant after the merger. The change in the composition of market entered is a strategic, equilibrium response due to the different configuration of network investment and competitors of different status in the markets.

Columns (4) to (6) of Table 6 report simulation results using estimates that do not account for endogeneity in 3G and neighboring 4G-LTE deployment (the first column in panel (2) of Table 5). Comparing column (1) to column (4), we can see a small underestimation of entry occurrences and a small overestimation of the underserved population, creating the impression that the impact of ignoring endogeneity in X_2 in the counterfactual simulation is negligible. The predicted changes in market entry under the second (merger) and third (merger + DOJ-mandated divestiture) scenarios in columns (5) and (6), however, are very different from those predicted in columns (2) and (3). We create Table 7 to compare the magnitudes of merger effects and remedy effects under these two models.

As Table 7 shows, if we ignore endogeneity in X_2, we will significantly underestimate the effect of the T-Mobile and Sprint merger on the reduction of entry and on the population affected by reduced entry. Specifically, the number of unserved markets will increase by 28.6% with the 4-to-3 merger under the endogenous X_2 model but only 7.9% if the endogeneity in X_2 is ignored in estimation and simulation. Underestimation of similar magnitude also happens for population affected by reduced entry. For example, the merger will lead to 23.4% increase in population underserved and 29.4% increase in rural population underserved under the endogenous X_2 model but only 6.5% and 8.6% in corresponding metrics under the exogenous X_2 model. The DISH remedy (merger + DOJ-mandated divestiture to DISH) effects predicted by the exogenous X_2 model fare better than the merger effects. Still, the exogenous X_2 model predicts improvement in all outcomes of interest, while the endogeneous X_2 model predicts trade-offs in these outcomes. The key message is that researchers would paint a much rosier picture of the consequences due to the merger and of DISH remedy if they were to use the model without considering X_2's endogeneity. Relying on such a biased prediction, policy makers would lean more toward approving the proposed merger.

6 Conclusion

We propose a new method for estimating discrete Bayesian games with endogenous covariates. The approach is flexible enough to accommodate endogeneity due to player- or game-level unobserved heterogeneity. We apply the method to estimate an entry game of 4G LTE deployments between major wireless service providers in the U.S. In this setting, existing 3G network investment and neighboring 4G-LTE investment are endogenous covariates. We find that a hypothetical merger between T-Mobile and Sprint would reduce 4G-LTE deployment significantly and that the divestiture remedy would not completely reverse the negative outcomes of the merger. More importantly, we show that incorporating the endogeneity of network investment affects our estimates of economic primitives, the counterfactual simulations under the hypothetical merger, and the policy implications. Both the merger and the remedy would lead to changes in the composition of markets entered and, in turn, different populations covered by cellphone services, and the effects captured by a model with endogenous network investment differs from those ignoring such endogeneity. Based on our results, we recommend that competition and regulatory authorities fully consider the multi-dimensional trade-offs between market power effects and efficiency gains from drastic changes in market structure due to mergers or entry and exit events.[43]

[43] This point is illustrated well by Bourreau, Yutec, and Verboven (2021), which shows that new entry in the French mobile telecommunications market led to a flurry of "fight brands" by incumbents.

References

Aguirregabiria, V. and P. Mira (2002). Swapping the nested fixed point algorithm: A class of estimators for discrete markov decision models. *Econometrica 70*(4), 1519–1543.

Aguirregabiria, V. and P. Mira (2007). Sequential estimation of dynamic discrete games. *Econometrica 75*(1), 1–53.

Aguirregabiria, V. and P. Mira (2019). Identification of games of incomplete information with multiple equilibria and unobserved heterogeneity. *Quantitative Economics 10*(4), 1659–1701.

Aradillas-Lopez, A. (2010). Semiparametric estimation of a simultaneous game with incomplete information. *Journal of Econometrics 157*(2), 409–431.

Aradillas-López, A. (2020). The econometrics of static games. *Annual Review of Economics 12*, 135–165.

Aradillas-Lopez, A. and A. Gandhi (2016). Estimation of games with ordered actions: An application to chain-store entry. *Quantitative Economics 7*(3), 727–780.

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010). Estimating static models of strategic interactions. *Journal of Business & Economic Statistics 28*(4), 469–482.

Berry, S. T. and J. Waldfogel (2001). Do mergers increase product variety? evidence from radio broadcasting. *The Quarterly Journal of Economics 116*(3), 1009–1025.

Blundell, R. W. and J. L. Powell (2004). Endogeneity in semiparametric binary response models. *The Review of Economic Studies 71*(3), 655–679.

Bourreau, M., S. Yutec, and F. Verboven (2021). Market entry, fighting brands and tacit collusion: Evidence from the french mobile telecommunications market. *Working paper, KU Leuven*.

Brock, W. A. and S. N. Durlauf (2001). Discrete choice with social interactions. *The Review of Economic Studies 5*(2), 235–260.

Caradonna, P., N. H. Miller, and G. Shue (2021). Mergers, entry, and consumer welfare. *Working paper, Georgetown University*.

Chesher, A. (2003). Identification in nonseparable models. *Econometrica 71*(5), 1405–1441.

Ciliberto, F., C. Murry, and E. Tamer (2021). Market structure and competition in airline markets. *Journal of Political Economy 120*(11).

Das, M., W. K. Newey, and F. Vella (2003). Nonparametric estimation of sample selection models. *The Review of Economic Studies 70*(1), 33–58.

De Paula, A. and X. Tang (2012). Inference of signs of interaction effects in simultaneous games with incomplete information. *Econometrica 80*(1), 143–172.

D'Haultfœuille, X. and P. Février (2015). Identification of nonseparable triangular models with discrete instruments. *Econometrica 83*(3), 1199–1210.

Dong, Y. and A. Lewbel (2015). A simple estimator for binary choice models with endogenous regressors. *Econometric Reviews 34*(1-2), 82–105.

Economides, N., J. E. Kwoka Jr, T. Philippon, R. Seamans, H. J. Singer, M. Steinbaum, and L. J. White (2019). Assessing doj's proposed remedy in sprint/t-mobile: Can ex ante competitive conditions in wireless markets be restored? *T-Mobile: Can Ex Ante Competitive Conditions in Wireless Markets Be Restored*, 19–14.

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

Elliott, J., G. V. Houngbonon, M. Ivaldi, and P. Scott (2021). Market structure, investment and technical efficiencies in mobile telecommunications. *Working paper, New York University*.

Fan, Y. (2013). Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review 103*(5), 1598–1628.

Fan, Y. and M. Xiao (2015). Competition and subsidies in the deregulated us local telephone industry. *The RAND Journal of Economics 46*(4), 751–776.

Fan, Y. and C. Yang (2020). Competition, product proliferation, and welfare: A study of the u.s. smartphone market. *American Economic Journal: Microeconomics 12*(2), 99–134.

Fan, Y. and C. Yang (2021). Merger, product variety and firm entry: The retail craft beer market in california. *Working paper, University of Michigan*.

Florens, J.-P., J. J. Heckman, C. Meghir, and E. Vytlacil (2008). Identification of treatment effects using control functions in models with continuous, endogenous treatment and heterogeneous effects. *Econometrica 76*(5), 1191–1206.

Florens, J.-P. and E. Sbaï (2010). Local identification in empirical games of incomplete information. *Econometric Theory 26*(6), 1638–1662.

Genakos, C., T. Valletti, and F. Verboven (2018). Evaluating market consolidation in mobile communications. *Economic Policy 33*(93), 45–100.

Glaeser, E. and J. A. Scheinkman (2003). Nonmarket interactions. *Advances in Economics and Econometrics 1*, 339–369.

Grieco, P. L. (2014). Discrete games with flexible information structures: An application to local grocery markets. *The RAND Journal of Economics 45*(2), 303–340.

Hahn, J. and G. Ridder (2011). Conditional moment restrictions and triangular simultaneous equations. *Review of Economics and Statistics 93*(2), 683–689.

Heckman, J. J. (1978). Dummy endogenous variables in a simultaneous equation system. *Econometrica 46*(4), 931–959.

Heckman, J. J. and R. Robb (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics 30*(1-2), 239–267.

Hoderlein, S. (2014). Endogenous semiparametric binary choice model with heteroskedasticity. *Working paper*.

Horst, U. and J. A. Scheinkman (2006). Equilibria in systems of social interactions. *Journal of Economic Theory 130*, 44–77.

Imbens, G. W. and W. K. Newey (2009). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica 77*(5), 1481–1512.

Jackson, M. O., Z. Lin, and N. Yu (2020). Adjusting for peer-influence in propensity scoring when estimating treatment effects. *Working paper*.

Kasahara, H. and K. Shimotsu (2012). Sequential estimation of structural models with a fixed point constraint. *Econometrica 80*(5), 2303–2319.

Kasy, M. (2011). Identification in triangular systems using control functions. *Econometric Theory 27*(3), 663–671.

Kavalar, A. (2014). *Empirical Studies of the Market for Broadband Personal Communications Service Spectrum in the US*. Ph. D. thesis, UCLA.

Klein, R. and F. Vella (2010). Estimating a class of triangular simultaneous equations models without exclusion restrictions. *Journal of Econometrics 154*(2), 154–164.

Lee, L., J. Li, and X. Lin (2014). Binary choice models with social network under heterogeneous rational expectations. *Review of Economics and Statistics 96*(3), 402–417.

Lee, S. (2007). Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics 141*(2), 1131–1158.

Lewbel, A. (2000). Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables. *Journal of Econometrics 97*(1), 145–177.

Lewbel, A. and X. Tang (2015). Identification and estimation of games with incomplete information using excluded regressors. *Journal of Econometrics 189*(1), 229–244.

Li, S., J. Mazur, Y. Park, J. Roberts, A. Sweeting, and J. Zhang (2019). Repositioning and market power after airline mergers. *working paper*.

Lin, Z., X. Tang, and N. N. Yu (2021). Uncovering heterogeneous social effects in binary choices. *Journal of Econometrics 222*(2), 957–973.

Lin, Z. and H. Xu (2017). Estimation of social-influence-dependent peer pressures in a large network game. *The Econometrics Journal 20*(3), 86–102.

Marcoux, M. (2020). Unobservable heterogeneity and identification in games: an application to mobile telecommunications. *Working paper, Université de Montréal*.

Misra, S. (2013). Markov chain monte carlo for incomplete information discrete games. *Quantitative Marketing and Economics 11*(1), 117–153.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Newey, W. K. (1987). Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics 36*(3), 231–250.

Newey, W. K., J. L. Powell, and F. Vella (1999). Nonparametric estimation of triangular simultaneous equations models. *Econometrica 67*(3), 565–603.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research 47*(1), 3–13.

Rivers, D. and Q. H. Vuong (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics 39*(3), 347–366.

Rothe, C. (2009). Semiparametric estimation of binary response models with endogenous regressors. *Journal of Econometrics 153*(1), 51–64.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica 55*(5), 999–1033.

Seim, K. (2006). An empirical model of firm entry with endogenous product-type choices. *The RAND Journal of Economics 37*(3), 619–640.

Soetevent, A. R., M. A. Haan, and P. Heijnen (2014). Do auctions and forced divestitures increase competition? evidence for retail gasoline markets. *The Journal of Industrial Economics 62*(3), 467–502.

Sweeting, A. (2009). The strategic timing incentives of commercial radio stations: An empirical analysis using multiple equilibria. *The RAND Journal of Economics 40*(4), 710–742.

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

Tang, X. (2010). Estimating simultaneous games with incomplete information under median restrictions. *Economics Letters 108*(3), 273–276.

Tenn, S. and J. M. Yun (2011). The success of divestitures in merger enforcement: Evidence from the j&j–pfizer transaction. *International Journal of Industrial Organization 29*(2), 273–282.

Todd, P. and K. I. Wolpin (2018). Accounting for mathematics performance of high school students in mexico: Estimating a coordination game in the classroom. *Journal of Political Economy 126*(6), 2608–2650.

Torgovitsky, A. (2015). Identification of nonseparable models using instruments with small support. *Econometrica 83*(3), 1185–1197.

Vytlacil, E. and N. Yildiz (2007). Dummy endogenous variables in weakly separable models. *Econometrica 75*(3), 757–779.

Wallsten, S. (2019). An economic analysis of the t-mobile - sprint merger. *Written Testimony Before the Subcommittee on Antitrust, Commercial, and Administrative Law, Committee on the Judiciary, United States House of Representatives*.

Wan, Y. and H. Xu (2014). Semiparametric identification of binary decision games of incomplete information with correlated private signals. *Journal of Econometrics 182*(2), 235–246.

Wollmann, T. G. (2018). Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review 108*(6), 1364–1406.

Xiao, M. and Z. Yuan (2021). License complementarity and package bidding: U.s. spectrum auctions. *Available at SSRN 3266253*.

Xu, H. (2018). Social interactions on large networks: a game theoretic approach. *International Economic Review 59*(1), 257–284.

Appendix A Proofs

Proof of Lemma 1. The existence of BNEs follows from the continuity of the mapping Γ and an application of the Schauder Fixed-point Theorem. We show the uniqueness of BNE by contradiction.

Suppose that there are two distinct BNEs, $P^1 \equiv (P_1^1, \cdots, P_K^1) \neq (P_1^2, \cdots, P_K^2) \equiv P^2$. From Equation (6), we have

$$\begin{aligned} |P_k^1 - P_k^2| &= \Big| f_k\Big(X_{k1}'\beta_k + X_{k2}'\gamma_k + \alpha_k \sum_{j\neq k} \tilde{P}_j + \sum_{j\in\mathcal{K}} V_j'\lambda_{k,j}\Big) \cdot \alpha_k \cdot \sum_{j\neq k}(P_j^1 - P_j^2)\Big| \\ &\leq \sup_t f_k(t) \cdot |\alpha_k| \cdot (K-1) \max_{j\in\mathcal{K}} |P_j^1 - P_j^2| < \max_{j\in\mathcal{K}} |P_j^1 - P_j^2|, \end{aligned}$$

where $\tilde{P} = (\tilde{P}_1, \cdots, \tilde{P}_K)$ is a vector of probabilities between P^1 and P^2. The first equality is due to the Mean Value theorem, and the second inequality is due to Assumption 5. Taking maximization of the left-hand side over $\mathcal{K}$ leads to a contradiction: $\max_{k\in\mathcal{K}} |P_k^1 - P_k^2| < \max_{j\in\mathcal{K}} |P_j^1 - P_j^2|$. □

Proof of Theorem 1. (Consistency of $\widehat{\theta}_{2SNPL}$) First, we show that $L_n(\cdot,\cdot;\widehat{\Pi}) \xrightarrow{p} L_0(\cdot,\cdot)$ uniformly over $\Theta \times P$. By the mean value theorem, for any $\theta \in \Theta$, $P \in \mathcal{P}$,

$$L_n(\theta, P; \widehat{\Pi}) - L_n(\theta, P; \Pi_0) = \nabla_\Pi L_n(\theta, P; \Pi^+)(\widehat{\Pi} - \Pi_0), \tag{11}$$

where Π^+ denotes an intermediate value between $\widehat{\Pi}$ and Π_0. By (11) and the triangular inequality,

$$\begin{aligned} &\sup_{\theta,P} \left| L_n(\theta, P; \widehat{\Pi}) - L_0(\theta, P) \right| \\ \leq\ &\sup_{\theta,P} |\nabla_\Pi L_n(\theta, P; \Pi^+)| \left| \widehat{\Pi} - \Pi_0 \right| + \sup_{\theta,P} |L_n(\theta, P; \Pi_0) - L_0(\theta, P)| \,. \end{aligned}$$

Under our maintained conditions, $\sup_{\theta,P} |\nabla_{\Pi} L_n(\theta, P; \Pi^+)| = O_p(1)$. Because $\widehat{\Pi} \xrightarrow{p} \Pi_0$, the first term on the right-hand side of the inequality is $o_p(1)$. By Assumption 4-(ii) and the fact that $l_i(\theta, P; \Pi_0)$ is continuous at each θ, P with probability one, the second term on the right-hand side of the inequality is $o_p(1)$. This establishes the uniform convergence of L_n to $L_0(\cdot)$ over $\Theta \times \mathcal{P}$.

Note that, by Assumption 3-(ii) and the Kullback-Leibler information inequality, (θ_0, P^0) uniquely maximizes $L_0(\theta, P)$ in the set Λ_0. Define

$$T(\theta, P; \Pi_0) \equiv \max_{c \in \Theta} \Big\{ L_0(c, P; \Pi_0) \Big\} - L_0(\theta, P; \Pi_0),$$

where we write out dependence of L_0 on Π_0 explicitly. Because $L_0(\theta, P; \Pi_0)$ is continuous and $\Theta \times \mathcal{P}$ is compact, Berge's maximum theorem establishes that $T(\theta, P; \Pi_0)$ is a continuous function. By construction, $T(\theta, P; \Pi_0) \geq 0$ for any (θ, P). Define

$$\mathcal{E} \equiv \Big\{ (\theta, P) \in \Theta \times \mathcal{P} : P = \Gamma(\theta, P; \Pi_0) \Big\}.$$

Since $\Theta \times \mathcal{P}$ is compact and Γ is continuous, $\mathcal{E}$ is a compact set. By definition, Λ_0 is a subset of $\mathcal{E}$. For each element in Λ_0, consider an arbitrarily small open ball that contains it. Let $B_\epsilon(\theta_0, P^0)$ denote the union of such open balls containing elements of Λ_0. Let B_ϵ^c denote the complement of B_ϵ. We then see that $B_\epsilon^c(\theta_0, P^0) \cap \mathcal{E}$ is also compact. Define the constant

$$\tau \equiv \min_{(\theta,P) \in B_\epsilon^c(\theta_0, P^0) \cap \mathcal{E}} T(\theta, P; \Pi_0) > 0. \tag{12}$$

Define the event

$$A_n \equiv \Big\{ \Big| L_n(\theta, P; \widehat{\Pi}) - L_0(\theta, P; \Pi_0) \Big| < \tau/2 \text{ for all } (\theta, P) \in \Theta \times \mathcal{P} \Big\}.$$

Let $\left(\theta_n^*, P_n^*\right)$ be an element of Λ_n. Then, A_n implies that

$$L_0\Big(\theta_n^*, P_n^*; \Pi_0\Big) > L_n\Big(\theta_n^*, P_n^*; \widehat{\Pi}\Big) - \frac{\tau}{2}; \text{ and}$$
$$L_n\Big(\theta, P_n^*; \widehat{\Pi}\Big) > L_0\Big(\theta, P_n^*; \Pi_0\Big) - \frac{\tau}{2} \text{ for any } \theta \in \Theta.$$

Besides, $L_n\Big(\theta_n^*, P_n^*; \widehat{\Pi}\Big) \geq L_n\Big(\theta, P_n^*; \widehat{\Pi}\Big)$ by definition of Λ_n. Thus,

$$L_0\Big(\theta_n^*, P_n^*; \Pi_0\Big) > L_n\Big(\theta_n^*, P_n^*; \widehat{\Pi}\Big) - \frac{\tau}{2} \geq L_n\Big(\theta, P_n^*; \widehat{\Pi}\Big) - \frac{\tau}{2} > L_0\Big(\theta, P_n^*; \Pi_0\Big) - \tau$$

for any $\theta \in \Theta$. Therefore,

$$\begin{aligned} A_n &\Rightarrow \left\{L_0\left(\theta_n^*, P_n^*; \Pi_0\right) > L_0\left(\theta, P_n^*; \Pi_0\right) - \tau, \forall \theta \in \Theta\right\}, \\ &\Rightarrow \left\{L_0\left(\theta_n^*, P_n^*; \Pi_0\right) > \max_{\theta \in \Theta} L_0\left(\theta, P_n^*; \Pi_0\right) - \tau\right\}, \\ &\Rightarrow \left\{\tau > T\left(\theta_n^*, P_n^*; \Pi_0\right)\right\}, \\ &\Rightarrow \left\{\min_{(\theta,P)\in B_\epsilon^c(\theta_0;P^0)\cap\mathcal{E}} T(\theta, P; \Pi_0) > T\left(\theta_n^*, P_n^*; \Pi_0\right)\right\} \text{ by (12)}, \\ &\Rightarrow \left\{(\theta_n^*, P_n^*) \in B_\epsilon(\theta_0, P^0)\right\}. \end{aligned}$$

The last induction uses the fact that $(\theta_n^*, P_n^*) \in \mathcal{E}$. Therefore, $\Pr(A_n) \leq \Pr\left((\theta_n^*, P_n^*) \in B_\epsilon(\theta_0; P^0)\right)$. By the uniform convergence of $L_n(\cdot; \widehat{\Pi})$ to $L_0(\cdot)$, $\Pr(A_n) \to 1$ as $n \to \infty$. Thus,

$$\Pr\left((\theta_n^*, P_n^*) \in B_\epsilon(\theta_0; P^0)\right) \to 1. \tag{13}$$

For the case in which Λ_0 is a singleton, this suffices for consistency of $\widehat{\theta}_{2SNPL}$.

In the general case in which Λ_0 has multiple elements, the proof follows from the same arguments in Aguirregabiria and Mira (2007), who proceed by showing the following results sequentially: (1) ϕ_n converges to ϕ_0 in probability uniformly in a neighborhood around P^0; (2) with probability approaching 1, there exists an element (θ_n^*, P_n^*) of Λ_n in any open ball around (θ_0, P^0); and (3) with probability approaching 1, the 2SNPL estimator is the element of Λ_n that belongs to an open ball around (θ_0, P^0).

(Asymptotic Normality of $\widehat{\theta}_{2SNPL}$) We now derive the limit distribution of $\widehat{\theta}_{2SNPL}$. To simplify notation, we drop the subscript $_{2SNPL}$ from the notation of this estimator in the proof below. By definition,

$$\frac{1}{n}\sum_{i=1}^{n} \nabla_\theta l_i(\hat{\theta}, \hat{P}; \widehat{\Pi}) = 0 \text{ and } \hat{P} - \Gamma(\hat{\theta}, \hat{P}; \widehat{\Pi}) = 0.$$

A stochastic mean value theorem between $(\theta_0, P^0; \Pi_0)$ and $(\hat{\theta}, \hat{P}; \widehat{\Pi})$, together with consistency of $(\hat{\theta}, \hat{P})$ and $\widehat{\Pi}$, implies that

$$\begin{aligned} \frac{1}{\sqrt{n}}\sum_{i=1}^{n} s_{\theta,i} - \Omega_{\theta\theta}\sqrt{n}\left(\hat{\theta} - \theta_0\right) - \Omega_{\theta P}\sqrt{n}\left(\hat{P} - P^0\right) - \Omega_{\theta\Pi}\sqrt{n}\left(\widehat{\Pi} - \Pi_0\right) &= o_p(1), \\ \left(I - \Gamma_P^0\right)\sqrt{n}(\hat{P} - P^0) - \Gamma_\theta^0\sqrt{n}(\hat{\theta} - \theta_0) - \Gamma_\Pi^0\sqrt{n}(\widehat{\Pi} - \Pi_0) &= o_p(1). \end{aligned}$$

Solving for $\sqrt{n}(\hat{P} - P^0)$ from the second set of equations and substituting into the first set, we get

$$
\begin{aligned}
&\underbrace{[\Omega_{\theta\theta} + \Omega_{\theta P}(I - \Gamma_P^0)^{-1}\Gamma_\theta^0]}_{\equiv M} \sqrt{n}(\hat{\theta} - \theta_0) \\
= \;& \frac{1}{\sqrt{n}} \sum_{i=1}^{n} s_{\theta,i} - [\Omega_{\theta P}(I - \Gamma_P^0)^{-1}\Gamma_\Pi^0 + \Omega_{\theta\Pi}] \sqrt{n}(\hat{\Pi} - \Pi_0) + o_p(1) \\
= \;& \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \underbrace{\left\{ s_{\theta,i} - [\Omega_{\theta P}(I - \Gamma_P^0)^{-1}\Gamma_\Pi^0 + \Omega_{\theta\Pi}] r_{0,i} \right\}}_{\equiv \tilde{s}_i} + o_p(1),
\end{aligned}
$$

where the second equality uses the asymptotic linear representation of $\sqrt{n}(\widehat{\Pi} - \Pi_0)$ and its influence function $r_{0,i} \equiv r_i(\Pi_0)$. The asymptotic distribution of $\hat{\theta}$ then follows from the continuous mapping theorem.

□

Appendix B Cellular Network Explained

A cellphone is a portable telephone that can make and receive calls over a radio frequency ("spectrum") while the user is moving within a service area. When a user makes a phone call or sends a message, her cellphone converts her voice or message into electrical signals, which are transmitted from her location to the nearest cell tower via radio waves. The network of cell towers then relays the radio waves to the receiver's cellphone, which converts it to electrical signals and then back to sound, text, or image again. In this process, data travel in a "cellular network," which is composed of cellphones, base transceiver stations ("cell sites"), mobile telephone switching offices, and the public switched telephone network. A cellphone is a type of Mobile Subscriber Unit, which consists of a control unit and a transceiver that transmits and receives radio transmissions to and from a cell site. The term cell site refers to the physical location of radio equipment that provides coverage within a cell. The hardware located at a cell site includes power sources, interface equipment, radio frequency transmitters and receivers, and antenna systems. A mobile telephone switching office is the central office for mobile switching. It houses the mobile switching center, field monitoring, and relay stations for switching calls from cell sites to wire-line central offices. The public switched telephone network is made up of local networks, exchange area networks, and the long-haul network that interconnect telephones and other communication devices on a worldwide basis. Boccuzzi(2019) describes the basics of cellular communications.

A new generation of network technology has arrived in almost every decade since the inception of such technology. The first two generations (0G and 1G) were before the widespread use of cellphones.[44] In the 1990s, 2G started the use of digital transmission

[44] 0G refers to pre-cellphone mobile technology, such as radio telephones that were placed in cars before the advent of cellphones. 1G refers to Analog Cellular Networks, which employ multiple cell sites to transfer calls from one site to the next as the user travels between cell sites during a conversation.

instead of analog transmission, marking the start of widespread use of cellphones in our lives. In the 2000s, 3G was the predominant technology, and in the 2010s, it was 4G-LTE. Now we are facing the transformation from 4G-LTE to 5G, the newest generation of network technology.

Appendix C Robustness of Table 5

We check the robustness of our Table 5's results by: 1) restricting our analysis to homogeneous competition effects; 2) using any generation of technology (2G, 3G, and 4G non-LTE combined) instead of just 3G to measure a firm's previous network investment in the focal market; 3) using a firm's 4G deployment in neighboring markets in 2015, instead of that in 2018, to measure the firm's network investment in neighboring markets.

Table 8 reports the results from these alternative specifications. In specification (1), the Big Four are treated as equal competitors. Results in this specification are very close to specification (1) in Table 5, suggesting only small differences in how AT&T/Verizon and T-Mobile/Sprint reacted to expected competition. In specification (2), for $X_{2,1}$, we expand from 3G to include any previous generation of technology deployment in the focal market. Compared with specification (1) in Table 5, the main change is that the estimate of the coefficient for V_1 and for $X_{2,1}$ become smaller in their absolute values. In specification (3), we do not consider the concurrent deployment of 4G-LTE in the neighboring census tracts; instead, we restrict 4G-LTE deployment in the neighboring census tracts to the status quo before the start of the entry game. Compared with specification (1) in Table 5, the main change is that the estimate of the coefficient for V_2 and for $X_{2,2}$ become smaller in their absolute values. The competition effects in specification (3) also become much smaller. In specifications (2) and (3) of Table 8, the qualitative results stay the same as Table 5.

In summary, these alternative specifications often produce marginally different magnitudes in estimates, but they all point to the importance of the network investment effect, as well as a consistently negative expected competition effect.

Table 8: Robustness of Table 5

	Homogeneous Competition		Use any G for $X_{2,1}$		Use 2015's 4G for $X_{2,2}$	
	(1)		(2)		(3)	
Variable	Est.	S.E.	Est.	S.E.	Est.	S.E.
Pop (in 000's)	0.077***	0.019	0.068***	0.019	0.051***	0.015
% Female	0.362	0.685	0.419	0.688	0.357	0.520
% Senior	-1.054 ***	0.387	-1.148***	0.402	-0.697**	0.338
% White	-0.416	0.481	-0.206	0.468	-0.725*	0.395
% Black	-0.524	0.511	-0.357	0.500	-1.000**	0.413
% Native	-1.546 ***	0.483	-1.360***	0.479	-1.374***	0.398
% Asian	0.931	1.334	1.452	1.253	0.060	1.036
% Hispanic	-0.526*	0.296	-0.560*	0.306	-0.505**	0.232
% College	-0.008	0393	0.034	0.389	0.235	0.320
% Labor force	-1.023***	0.286	-0.791***	0.289	-1.184***	0.243
% Work home	0.236	0.524	0.159	0.559	-0.015	0.430
% Long commute	0.039	0.283	-0.063	0.294	0.562***	0.217
HH income	-0.002	0.003	-0.004	0.003	0.001	0.002
HH size	-0.044*	0.023	-0.037	0.023	-0.018	0.019
Pop density	-0.125*	0.074	-0.126	0.078	-0.063	0.070
% Rural	0.239***	0.092	0.203**	0.097	0.080	0.074
Mostly Water	-1.734 ***	0.485	-1.740***	0.485	-1.578***	0.410
# Incumbents	-0.126***	0.040	-0.115***	0.042	-0.039	0.034
Intercept	0.998	0.695	0.717	0.684	1.240**	0.540
Network Investment Effects						
$X_{2,1}$	3.858***	0.251	3.289***	0.222	3.498***	0.235
$X_{2,2}$	1.639 ***	0.143	1.630***	0.142	0.240**	0.119
V_1	-1.986***	0.218	-1.453***	0.210	-1.086***	0.183
V_2	2.005***	0.210	2.001***	0.206	0.814***	0.204
Expected Competition Effects						
Any P.E.	-0.994 ***	0.061	-	-	-	-
Strong P.E.	-	-	-0.867***	0.075	-0.633***	0.084
Weak P.E.	-	-	-0.940***	0.063	-0.533***	0.064

Notes: The results are based on 6,244 observations, corresponding to two to four potential entrants for 2,582 Census tracts. Standard errors are obtained from resampling of markets with replacement 1,000 times). Asterisks indicate the statistical significance at the 1% (***), 5% (**), and 10% (*) levels.