Endogeneity in Games with Incomplete Information: U.S. Cellphone Service Deployment*

Zhongjian Lin[†] University of Georgia

Xun Tang[‡] Rice University

Mo Xiao[§] University of Arizona

May 10, 2024

Abstract

In some discrete games with incomplete information, payoff-relevant states are influenced by unobserved heterogeneity that directly affects strategic decisions. When ignored, such endogeneity potentially leads to problematic parameter inference and policy implications. We introduce a control-function (CF) approach for estimating such games, and apply the method to an entry game of deploying 4G-LTE technology between major U.S. cellphone service providers. Unlike CF methods in single-agent contexts, our CF approach in the context of Bayesian games is based on new conditions on how unobserved *market* and player heterogeneity correlate with sources of endogeneity and instruments. Taking network investment as endogenous, we find that a hypothetical T-Mobile and Sprint merger would reduce 4G-LTE deployment across the local markets in our sample, and disproportionately decrease rural coverage. Ignoring such endogeneity would under-predict the negative impacts of the merger, therefore favoring its approval.

Keywords: Endogeneity, Discrete Games with incomplete information, 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, Paul Ellickson, Paul Grieco, Keisuke Hirano, Krzysztof Karbownik, Ruixuan Liu, Essie Maasoumi, Mathieu Marcoux, Sue Mialon, and Joris Pinkse for useful comments and suggestions. All remaining errors are our own.

[†]John Munro Godfrey, Sr. Department of Economics, Terry College of Business, University of Georgia, Athens, GA, 30602. Email: zhongjian.lin@uga.edu

[‡]Department of Economics, Rice University, Houston, TX 77005. Email: xun.tang@rice.edu

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

1 Introduction

We propose a control-function approach for estimating discrete games with incomplete information (a.k.a. Bayesian games) when observable payoff-relevant covariates are endogenous. Games with incomplete information provide a powerful framework for analyzing strategic interaction among individuals or firms with private information (e.g. types or signals) and have been studied in a wide range of applied contexts. Examples include location choices in the video retail industry in Seim (2006); timing of commercials by radio stations in Sweeting (2009); choices of movie release dates in Einav (2010); market entry and exit of grocery stores in Grieco (2014); choices of effort by students and teachers in classrooms in Todd and Wolpin (2018); and choices of fitness exercises by adolescents in Jackson, Lin, and Yu (2022). An important assumption required for inference in these empirical studies is that the covariates are exogenous at both the player- and the game-level.

When covariates in a game with incomplete information are endogenous, identification and estimation require further assumptions on the joint distribution of covariates, private types, and instruments. This endogeneity issue poses challenges for model inference that are analogous to endogeneity in single-agent qualitative response models, but are aggravated under strategic interaction. Specifically, dealing with endogenous covariates in individual choices in such cases requires constructing a complete vector of individual-specific control functions for *all* players. 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 because players' equilibrium beliefs need to be conditioned on private types.[1]

Endogeneity in covariates is common in environments with strategic interaction. Berry and Reiss (2007) pointed out that the issues of "endogenous scale of operations" and "endogenous product characteristics in a continuous space" are "empirically important and valuable extensions of the existing empirical literature on structural

[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/game-level factor. 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 $E(D_j|X_i, X_j, u_i)$.

models for market structure."[2] For a specific 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 (which measures the existing network scale) and 4G-LTE deployment in neighboring markets (which measures the spatial network spillover) are both important covariates that influence its decision to enter a local 4G-LTE 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 the investment cost of the 3G and 4G deployment in the focal and neighboring markets. For another example, a provider's spectrum holdings can be strongly correlated between focal and neighboring markets to affect deployment decisions across these markets.[3] More broadly speaking, endogeneity is a concern in social-economic settings in which some payoff-relevant states involve past decisions (e.g., advertising expenditure, R&D expenses) that were influenced by latent, unreported player- or game-level factors. We are not aware of any paper that allows for such flexible sources of endogeneity in the empirical analysis of discrete games with incomplete information 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 discrete games with incomplete information in two ways. First, we introduce a general, feasible control-function method for estimating binary games with incomplete information in the presence of 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.[4] Specifically, the first

[2] Dunne, Klimek, Roberts, and Xu (2013) echo this point, stating "... (a direction for future research) involves incorporating firm-level heterogeneity in profits, fixed costs, and/or entry costs that is correlated over time for individual firms." Firm-level covariates that display persistence over time are very likely endogenous.

[3] Spectrum licenses are initially allocated by the Federal Communication Commission's auctions. Due to frequent activities in the secondary market, each firm's spectrum holdings are imperfectly observed by other firms and unobserved by the researchers.

[4] Grieco (2014) studies discrete games with exogenous covariates but a flexible information structure

(structural, outcome) equation in the system specifies how firm profits are affected by competition and firm-level endogenous states. We model these effects as policy-invariant parameters, and use them to predict market outcomes under counterfactual policy scenarios. The second (auxiliary, selection/endogeneity) equation is reduced-form in nature; it relates the endogenous states directly to profit-relevant exogenous covariates as well as instruments (which affect firm profits only indirectly through the endogenous states).

Our control function approach for dealing with endogenous covariates in Bayesian games should *not* be viewed as a direct extension of CF methods in single-agent contexts. This is because in such games the endogenous covariates enter the players' strategies non-linearly through information sets, on which individual beliefs are conditioned. Consequently, our CF approach in such games is based on new, distinctive conditions that could accommodate a rather flexible correlation between unobserved *market-level* and player-level heterogeneity and the endogenous payoff states. We elaborate on the economic substance of these identifying conditions in the text, following their introduction (as Assumption 1) in Section 2.

A notable advantage of this model is that it allows us to identify the effects of competition and endogenous states on market outcomes and to conduct counterfactual policy analysis without having to structurally model another layer of strategic interactions which (pre-)determine the endogenous states. Such advantage exists because of two aspects: (a) the reduced-form endogeneity equation is flexible enough to capture realistic sources of endogeneity, and incorporates instruments for valid inference despite endogeneity; (b) in our policy analysis we consider immediate post-merger scenarios in which the status of the new merger's endogenous states (network deployment) is completely specified. This latter aspect means we do not need post-merger simulation of endogenous states, thus dispensing with the need for a structural module about how endogenous states are pre-determined in a "first-stage" game. More generally, the CF method is suitable for analyzing the implication of policies that do not perturb the relation between endogenous states and instruments in (2).

The CF method we propose also has clear advantages over an alternative,

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 (2022) recovers the unobserved heterogeneity of competitors in the Canadian telecommunications industry, using the reduced form of their bids for spectrum licenses. He then uses the variation of such heterogeneity to identify incumbents' responses to the new entrant's decisions. In his case, the bids for spectrum licenses do not enter the network investment game directly as payoff-relevant states. In contrast, we construct control functions from endogenous, payoff-relevant states and exogenous instruments.

fully parametric approach which would require researchers to completely specify the distribution of u_k conditional on *all* endogenous covariates X_{k2}. First, our method is robust and intuitive, because the stochastic restrictions we maintain on the unobserved errors are not only nonparametric, but also have direct structural, economic interpretation. Second, our method is suitable for showing robust model identification. In comparison, a fully parametric approach would count on the assumed form of likelihood to attain identification (which in many cases could only be done locally by checking the singularity of the Jacobian of the likelihood).

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 (in Appendix B) 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,[5] the control function approach has proliferated due to its simplicity, flexibility and wide applicability.[6] For example, Petrin and Train (2010) used control functions to deal with endogenous product characteristics (such as prices) in a consumer demand model, using marginal cost shifters as instruments. We contribute to this extensive literature by bringing the control function approach into a setting of discrete games with incomplete information.[7] We combine control functions with a nested pseudo likelihood method to handle the simultaneity embedded in a game with incomplete

[5]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 outcomes and endogenous variables can be identified using discrete instruments.

[6]Since its inception by Heckman and Robb (1985), the control function approach has been used in various 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), Hahn and Ridder (2011), and Kasy (2011) among others.

[7]For econometric analyses of static games with incomplete information, 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).

information and the endogeneity in regressors at the same time.[8]

Aguirregabiria and Mira (2019) showed the identification of Bayesian games when unobserved heterogeneity may also affect the selection between multiple equilibria, thus leading to another source of endogeneity. In comparison, our CF method focuses on endogenous covariates in a data-generating process without multiple equilibria, and as such cannot be applied to deal with endogeneity in equilibrium selection. On the other hand, unlike Aguirregabiria and Mira (2019), our CF method does not restrict the support of the unobserved heterogeneity to be finite. Another major distinction is that Aguirregabiria and Mira (2019) used methods from the finite mixture literature to recover the players' choice probabilities conditional on the unobserved heterogeneity as an necessary step to identify players' payoffs. In contrast, our CF method does not require such a step, and can be used to estimate the payoff parameters in a simple two-step procedure.

Our second contribution is to provide new empirical insights in a setting where endogeneity in firm states affects the inference of competition effect. We apply the 2SNPL estimator to analyze a hypothetical 2016 T-Mobile and Sprint merger on cellphone technology deployment in a selected sample of mostly rural markets.[9] During 2016 and 2018, AT&T, Verizon, T-Mobile and Sprint were the main competitors in the U.S cellphone service industry. Having rolled out 4G-LTE technology across the states, these firms were now strategically considering whether to enter the remaining scattered markets left open to 4G-LTE deployment. The 4-to-3 merger is clearly a mover-and-shaker event in this industry, which could substantially change firm strategies of entering local markets. Our sample consists of isolated markets that had not deployed the latest generation of cellphone technology at the time. The geographically dispersed, isolated markets allow us to treat each market as an independent entry game, so we can focus on the first-order question about the relation between endogenous states and entry decisions on focal markets.[10]

Given the scope of our sample, our goal is not to assess the overall merger effect on the industry level. Rather, we aim at providing insights to antitrust and

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

[9] 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.

[10] A full account of large scale, interdependent entry of 4G-LTE entry across the states would require modeling the providers' strategic optimization of deployment on a national level.

regulatory agencies who care about the possibility that merger exacerbates the urban-rural digital divide. 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 hypothetical 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.[11] In the case of the 2020 T-Mobile/Sprint merger decision, 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 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 in our sample, despite the merged firm becoming a strong competitor and owning 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.[12] Moreover, our simulations show that the addition of a fourth

[11]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.

[12]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

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 in our sample. 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 overestimate the total number of entry occurrences but underestimate 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 endogeneity of assets and divestiture.[13] 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 games with incomplete information.

The paper is organized as follows. Section 2 introduces the discrete games with incomplete information and endogenous states and characterizes the Bayesian Nash equilibrium. Section 3 describes the two-step nested pseudo likelihood estimator (2SNPL) and derives its asymptotic properties. Section 4 studies the 4G-LTE entry game of AT&T, Verizon, T-Mobile and Sprint, comparing model estimates and policy

did not change significantly.

[13]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.

implications with and without accounting for endogenous covariates. Section 5 concludes. All proofs, technical details, and robustness checks are rendered in the appendices.

2 Discrete Bayesian Games with Endogenous States

2.1 The Model and Equilibrium

Consider a game of simultaneous discrete choices with incomplete information 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, which consists of a $d_1 \times 1$ vector of exogenous covariates X_{k1}, and a $d_2 \times 1$ vector of endogenous covariates X_{k2}. Let Z_k be a $d_z \times 1$ vector of instruments, with $d_z \geq d_2$. Each player k observes a payoff shock, a.k.a. structural error, $u_k \in \mathbb{R}$, and makes a simultaneous entry decision $Y_k \in \{0, 1\}$ based on the public information $\mathbb{I} \equiv \{X_k, Z_k\}_{k \in \mathcal{K}}$ and its individual 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. Payoff specification similar to Equation (1) is common in the literature of empirical discrete games (see Berry (1992), Seim (2006), and Jia (2008) for examples).[14] The structural errors u_k's absorb all factors that affect firms' *ex post* payoffs but are not reported in the data. 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 $d_2 \times (d_1 + d_z)$ matrix of constant coefficients, and V_k contains player- and/or game-level unobserved heterogeneity that affect the determination of X_{k2}. The regressor X_{k2} in Equation (1) is endogenous when $V_k \in \mathbb{R}^{d_2}$ and $u_k \in \mathbb{R}$ are correlated. Instrument validity requires the coefficients for Z_k in Π_k be non-zero.

The linearity of the auxiliary equation (2) is not essential for our method. 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 functions (CFs) V_k satisfy Assumption 1 below. In Section 4, we use a quadratic

[14]This tractable functional form is appealing, because it serves as a practical approximation of the expected discounted value of a firm's action. It is infeasible to construct a full-fledged engineering model to investigate the action's effect on a firm's long-run profits. Besides, lack of data on firm-level prices and quantities restricts the researchers' ability to write down a more primitive payoff function.

function of covariates in the first stage.

It is worth mentioning that our method also applies when the right-hand side of equation (2) involves Z_j for $j \neq k$, i.e., instruments specific to *other* players. In that case, V_k is recovered as residuals by regressing X_{k2} on X_{k1} and $(Z_j)_{j\in\mathcal{K}}$. Such a case is relevant, for example, if the endogenous covariates $(X_{k2})_{k\in\mathcal{K}}$ are jointly determined through player interactions even before the game, and can be represented or approximated through a system of simultaneous equations.

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.)

2.2 Identifying Assumption and Discussion

Our method for dealing with endogenous covariates in this model applies under intuitive conditions on the unobserved errors, which we formalize as follows. Assume:

$$u_k = g_k(V; \lambda_k) + \eta_k \text{ for } k = 1, 2, ..., K, \tag{4}$$

where $V \equiv \{V_j\}_{j\leq K}$, and $g_k(\cdot)$ is a *link function* characterized by a parameter vector λ_k.

Note that in our model, u_k, the structural payoff shock in player k's information set, should *not* be interpreted as the player's private information. Instead, the idiosyncratic error η_k in (4) represents *de facto* private information for player k after accounting for the vectors in V, which can be recovered from the common information of (X, Z) using (2). In this sense, V's are effectively public information for all players.

Assumption 1. *(i) $\{u_k, V_k\}_{k\leq K}$ are independent of $X_1 = \{X_{k1}\}_{k\leq K}$ and $Z = \{Z_k\}_{k\leq K}$ with zero means. (ii) Equation (4) holds with $\eta \equiv \{\eta_k\}_{k\leq K}$ independent from V, and η_k independent across the players.*

Part (i) of this assumption posits instrument exogeneity; the zero means of u_k and V_k are necessary location normalization when an intercept is included in the exogenous covariates X_{k1}. Part (ii) posits that the residual private information, after accounting for unobserved heterogeneity in V in the structural payoff shocks u_k, are independent from such unobserved heterogeneity. This condition ensures V serves as valid control functions, despite correlation between players' payoff shocks and strategic interaction in equilibrium. Part (ii) also requires the residual private information to be independent across individual players. We will show how Assumption 1 can be satisfied in our model under more primitive conditions below.

It is important to note that Assumption 1 allows the error vectors $\{u_k, V_k\}_{k\leq K}$ to be correlated across players through the control functions $\{V_k\}_{k\leq K}$. In particular, this allows V_k to be correlated through payoff-relevant information (such as market shocks or heterogeneity) that is commonly observed to all players but not reported in the data. This differs from single-agent models where observations of individual decision makers are considered independently and identically distributed.

Assumption 1-(ii) is flexible enough to accommodate different forms of endogeneity in X_{k2}, including those due to player-level or game-level unobserved heterogeneity. For example, consider a data-generating process where V_j is arbitrarily correlated across players $j \leq K$, through *game-level* unobserved heterogeneity. Suppose for each player k, u_k is a polynomial function of $\{V_j\}_{j\leq K}$ (with a known degree) plus idiosyncratic noises η_k, which are independent across the players and jointly independent of $\{X_1, Z, V\}$. Then Assumption 1 hold with λ_k in (4) being the polynomial coefficients.

Assumption 1-(ii) also allows for endogeneity caused by unreported *individual* heterogeneity, even under more restrictive special cases of *linear* link functions:

$$g_k(V; \lambda_k) = \sum_{j\leq K} V_j' \lambda_{k,j}, \tag{5}$$

where $\lambda_k \equiv \{\lambda_{k,j}\}_{j\leq K}$ are constant coefficients. For example, suppose $\{u_k, V_k\}_{k\leq K}$ are independent across the players. If, for each player k, the vector $\{u_k, V_k\}$ is multivariate normal with non-zero correlation between u_k and V_k (e.g., because of correlation through unobserved characteristics of player k), then Assumption 1 is satisfied with $\lambda_{k,k}$ being coefficients in a linear projection of u_k onto V_k, and $\lambda_{k,j} = 0$ for $j \neq k$.

More generally, Assumption 1-(ii) can also hold with a link function in (5) when (u_k, V_k) are correlated among the players. To see this, suppose u, V are jointly distributed as multivariate normal, with covariances Σ_u and Σ_V and cross-covariance Σ_{Vu}. Assumption 1-(ii) holds with a linear link function in (5) if $\Sigma_u - (\Sigma_{Vu})'(\Sigma_V)^{-1}\Sigma_{Vu}$ is diagonal. With $u \in \mathbb{R}^K$ and $V \in \mathbb{R}^L$, this amounts to $K(K-1)/2$ equality constraints on $K(K-1)/2 + L(L-1)/2 + KL$ free parameters in $\Sigma_u, \Sigma_V, \Sigma_{Vu}$.

Some earlier works dealt with market-level heterogeneity in discrete games by modeling how they enter in the payoffs and equilibria selection mechanism (see

Aguirregabiria and Mira, 2007, Arcidiacono and Miller, 2011, Collard-Wexler, 2013, Igami and Yang, 2016, Aguirregabiria and Mira, 2019). Our approach differs in that we accommodate such heterogeneity in the structural errors of players' payoffs without imposing further restrictions (other than Assumption 1).

2.3 Identification via Control Functions

For the rest of the paper, we present identification and estimation results for the case where the Assumption 1 holds with a linear link function as in (5). As mentioned in Section 2.2, this can be justified in certain settings where the payoff shock u and the control functions V are multivariate normal. Generalization to known, flexible forms in (4) would not pose any new conceptual challenge, and is left for future investigation.

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

$$Y_k = 1\{Y_k^* > 0\} = 1\Big\{X_{k1}'\beta_k + X_{k2}'\gamma_k + \alpha_k \sum_{j\neq k} \mathbb{E}_k(Y_j|\mathbb{I}) + \sum_{j\leq K} V_j'\lambda_{k,j} + \eta_k > 0\Big\}. \tag{6}$$

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 players in Assumption 1-(ii) implies that the equilibrium belief $E_k(Y_j|\mathbb{I}, u_k) = E_k(Y_j|\mathbb{I})$, and does not depend on the residual private information $\{\eta_j\}_{j\neq k}$.[15]

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

$$P = \Gamma(\theta, P), \tag{7}$$

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

$$\Gamma_k(\theta_k, P) \equiv 1 - F_k\Big(-X_{k1}'\beta_k - X_{k2}'\gamma_k - \alpha_k \sum_{j\neq k} P_j - \sum_{j\leq K} V_j'\lambda_{k,j}\Big), \tag{8}$$

and $\theta \equiv \{\theta_k\}_{k\leq K}$ with $\theta_k \equiv (\gamma_k', \beta_k', \alpha_k', \lambda_k')$. The model admits a *unique* psBNE under the following condition.

Assumption 2. *For each k, $|\alpha_k| < \frac{1}{(K-1)|\sup_t f_k(t)|}$.*

Our estimator works as long as the sample data is generated from a single equilibrium. Assumption 2 restricts the strength of interaction between players so

[15] Conditional on $\mathbb{I}$, the values of $\{V_k\}_{k\leq K}$ are fixed. Hence "conditioning on $\mathbb{I}$ and u_k" is equivalent to "conditioning on $\mathbb{I}$ and η_k." The claim then follows from the independence conditions in Assumption 1.

that Γ is a contraction mapping and the uniqueness of psBNE is guaranteed. Such uniqueness is crucial for counterfactual simulation.[16]

Let P^* denote the profile of conditional choice probabilities in a Bayesian Nash equilibrium. Identification using Equations (7) and (8) 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: there are more instruments in Z_k than endogenous covariates in X_{k2}.

The aforementioned rank condition for recovering payoff parameters holds generically, even with the instruments being on the market level (i.e., $Z_k = Z_0$ does not vary across players k) *and* having small support (e.g., $Z_0 \in \{0, 1\}$). To see this, note that $\{X_{k1}, V_1, V_2..., V_K\}$ are all exogenous variables, that X_{k2} is a function of X_{k1} and the instrument Z_0, and that expected competition $\sum_{j\neq k} P_j^*$ is determined in equilibrium as a non-linear function of a *larger* vector including *other* players' exogenous variables $\{X_{j1}, V_j\}_{j\leq K}$ as well as Z_0. Because the joint variation of $\{X_{j1}, V_j\}_{j\leq K}$ and Z_0 is not confined to any linear subspace, and because of the non-linearity of $\sum_{j\neq k} P_j^*$ in these arguments, the rank condition is expected to hold generically.[17]

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.

2.4 Further Discussion: Extensions

We conclude this section with remarks about extensions of our method. First, the identification of this model using our control function (CF) approach does not require parametric assumptions on the distribution of η_k. In principle, one can

[16]Assumption 2 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, 2022, Lin, Tang, and Yu, 2021) for the uniqueness of Bayesian Nash equilibrium.

[17]With $Z_k = Z_0$ not varying across k, and with d_z as small as possible (i.e., $d_z = d_2$), a vector that consists of the endogenous states $\{X_{j2}\}_{j\leq K}$ for *all* players and a full set of all exogenous variables $\{\{X_{j1}\}_{j\leq K}, Z_0\}$ is linearly dependent. This is because in such cases X_{k2} can be expressed as a linear function of other $X_{j2}, j \neq k$ and the full set of all exogenous variables. However, this does not pose an issue to the rank condition for identification mentioned above, because the vector $(X_{k1}', X_{k2}', \sum_{j\neq k} P_j^*, V_1', \cdots, V_K')$ only includes X_{k2} and X_{k1} for a *single* player k while all other $\{X_{j1}, X_{j2}\}_{j\neq k}$ are conditioned on in P_j^* non-linearly.

apply the semiparametric estimators in Aradillas-Lopez (2010) to the reduced form in equation (6) after plugging in estimated control functions V_k as additional covariates. However, the derivation of the asymptotic property of such a semiparametric estimator that accounts for the first-step estimation of CFs is beyond the scope of this paper. Instead, we choose to focus on a tractable version with a parametric distribution of η_k in order to highlight the merits of the CF method mentioned earlier.

Second, we can extend the CF method to similar games with incomplete information where the endogenous states are discrete rather than continuous. The most intuitive and feasible way to do so is by introducing parametric assumptions (such as multivariate normal) on the joint distribution of (u_k, V_k). In this case the control function variables take the form of the correction terms in Heckman (1978). Recently, Lin and Tang (2022) applied CFs for estimating a social interactions model where all group peers receive an endogenous binary treatment. Gu, Li, Lin, and Tang (2022) apply CFs to identify peer effects in social interactions models with endogenous sample selection. They also generalize the CF method to allow for correlated group fixed effects. The structural form of outcomes in those social interactions models is qualitatively different from the discrete games with incomplete information in this paper. Nevertheless, with appropriate changes, the CF approaches in those papers can be adjusted for use in the current context of discrete games.

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 the 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* $(x, z) \in \mathcal{X}^K \times \mathcal{Z}^K$. *(iv)* $\eta_{k,i}$ *are i.i.d. standard normal.*

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). Assumption 3(iv) adopts a normal parametrization of the distribution of $\eta_{k,i}$. The zero mean of $\eta_{k,i}$ is implied by conditions on (u_k, V_k) in Assumption 1(i); the unit variance of $\eta_{k,i}$ is a necessary scale normalization for estimation. In principle, one may use a different parametrization of the distribution of $\eta_{k,i}$, e.g., standard logistic.

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 $\Pi \equiv \{\Pi_k\}_{k \le K}$ and $f_{k,i}(\theta, P; \Pi)$ defined as

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

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.[18]

With a slight abuse of notation, we let $\Gamma(\theta, P; \Pi)$ denote the mapping $\Gamma(\theta, P)$ as defined in Equation (8) 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 $k \le K$. In the second stage, plug $\widehat{\Pi} \equiv \{\widehat{\Pi}_k\}_{k \le K}$ 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

[18] The term "pseudo likelihood" is used because the argument P in L_n may be an arbitrary profile of choice probabilities, not necessarily the equilibrium choice probabilities P^0.

initial guess from a reduced-form probit regression without endogenous covariates and strategic interactions terms.

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

$$\widehat{\theta}_s = \arg\max_{\theta\in\Theta} L_n(\theta, \widehat{P}_{s-1}; \widehat{\Pi}), \tag{9}$$

and update the choice probabilities recursively as

$$\widehat{P}_s = \Gamma(\widehat{\theta}_s, \widehat{P}_{s-1}; \widehat{\Pi}). \tag{10}$$

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 (9) and (10):

$$\phi_n(P) \equiv \Gamma(\tilde{\theta}_n(P), P; \widehat{\Pi}), \text{ where } \tilde{\theta}_n(P) \equiv \arg\max_{\theta\in\Theta} L_n(\theta, P; \widehat{\Pi}). \tag{11}$$

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.

For discrete games with incomplete information but no endogenous covariates, the econometrics literature proposed two-step estimators, with the first step involving nonparametric (e.g., kernel-based) estimates of the players' conditional choice probabilities (CCPs) in equilibrium. Examples include Aradillas-Lopez (2010) and Bajari et al. (2010). In principle, it is possible to adjust those methods for our case by incorporating the control functions. However, in our empirical context, kernel estimation of CCPs in the first step would be costly due to the curse of dimensionality (i.e. the number of covariates conditioned on in CCPs is large). In comparison, the iterative algorithm in our 2SNPL estimator can be applied using other flexible initial estimators of CCPs (e.g. flexible logit with a polynomial series of covariates) under the conditions in Theorem 1 and Theorem 2. This reduces the computation costs, and is shown to achieve fast convergence and stability in practice. We hope the 2SNPL estimator could complement existing methods for estimating discrete games with incomplete information in the same way the nested pseudo likelihood estimator in Aguirregabiria and Mira (2007) complements CCP-based methods in Hotz and Miller (1993) for dynamic models.

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

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

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

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 $(0,1)^{K\cdot|\mathcal{X}|\cdot|\mathcal{Z}|}$; (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 to 4, $\hat{\theta}_{2SNPL}$ 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

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

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.

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

Theorem 2. *Suppose that Assumptions 1 to 4 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 from Assumption 2 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.[19] 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 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

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

firms compete through strategic 4G-LTE deployment decisions in local markets. We model this deployment decision as an entry game with incomplete information because each firm observes its firm-and-market-specific shocks in the deployment payoff but not necessarily those shocks to its rivals. Other papers using similar information structures for investigating firm entry decisions in different contexts include Seim (2006) and Sweeting (2009). We model the competition between 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").[20] The time period we study 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 in a specific geographic area before providing services to consumers in that area. 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.[21] A firm with 3G deployment in a local market can repurpose the spectrum used by 3G to support 4G-LTE and can utilize existing facilities, such as cell towers, with upgraded equipment. Such investment also involves some 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.

In the following subsections, we describe the background of the U.S. cellphone

[20] 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.

[21] 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.

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.

4.1 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.[22]

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.

4.2 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.[23] 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.

Building a cellular network takes decades of physical and financial investment from

[22]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.

[23]We explain the components and evolution history of cellular networks in Appendix C.

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.[24] 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.[25] 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.

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

[25]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 as selling off 700 MHz, which was used mainly for 4G-LTE deployment.

4.3 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.[26] 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 roll-out of 5G services.[27]

On July 26, 2019, the DOJ approved the merger after T-Mobile and Sprint reached an agreement to sell Sprint's branded prepaid business,[28] 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.[29] 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).

[26]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).

[27]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.

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

[29]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 on a national level 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 inequality 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.[30] The merged firm, aided by "the unique combination of spectrum, sites and equipment of T-Mobile and Sprint",[31] 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 on local markets that had not been served by most national providers by then. 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 games with incomplete information to analyze the impact of the hypothetical merger and remedy, taking into account the firms' post-merger network consolidation and strategic responses.

4.4 Data Sources

We use three publicly available data sets to construct our sample. The first is the FCC's Mobile Deployment Form 477 Data from 2015 to 2018, which reports semi-annually each provider's 2G-4G coverage in every U.S. census block.[32] The FCC requires all

[30]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.

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

[32]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

facilities-based broadband providers to file Form 477, which discloses 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,[33] and 4G-LTE) by each provider, using a computationally intensive process.[34] 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 aggregate 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 the same variables as the ones we obtain from the 2016 American Community Survey 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.

4.5 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.

census block by each provider. Much of the information presented in the data description is based on the FCC's Public Notice (DA 16-1107), released on September 30, 2016.

[33]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.

[34]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.

4.5.1 Local markets: census tracts

Our first task is to define a local market. On the demand side, we take advantage of the immobile demand of the cellular services — a typical consumer would only order services from providers which offer service in their neighborhood. But how large is exactly a neighborhood? A census block, a census tract, or an entire county? We resort to the supply side to refine our market definition. In this study we investigate the marginal decision of a provider to deploy 4G-LTE to additional market. This is a local investment decision, not an overall entry/exit decision. The definition of a local market then boils down to the question that at what geographical level sunk costs of serving an area are committed.

A census block is the smallest geographic unit in the U.S. Census, amounting to more than 11 million 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 consider a larger market concept: the 73,057 census tracts in the U.S.. 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.[35] 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.[36] 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.[37] A county is simply too large and too heterogeneous for the definition of a local market. Based on the above comparison, we define census tracts as geographic markets based on which cellphone providers make investment and network deployment decisions.

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. 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

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

[36] 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.

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

between 0% and 10% coverage and a major peak at 100%.[38] 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.

4.5.2 Sample construction

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, mostly isolated, open markets. We focus on a potential entrant's decisions to enter these local, isolated markets.[39] 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 of interdependent local markets. 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, but had operated in at least one census tract in that state. A potential entrant is observed as having decided to enter a market if it made 4G-LTE deployment by the end of 2018. Outside of the Big Four, we do not consider other

[38]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.

[39]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.

providers as potential entrants in our game, but we count their presence in a census tract by the end of 2015 as part of the incumbent competition. Our sample consists of 2,582 census tracts that have at least two potential entrants by the end of 2015. Each census tract has an entry game, with two to four potential entrants. When there are only two potential entrants, it must be that the other two Big Four providers have already entered and are counted as incumbents. Still, most of the incumbents in a census tract, if there is any, are small, local or regional providers.

4.5.3 Summary statistics: the Big Four's cellphone deployment

In Table 1, 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 1 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 (denoted by $X_{k2,1}$). The second is the firm's 4G-LTE deployment in neighboring tracts by the end of 2018 (denoted by $X_{k2,2}$). 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. Specifically, $X_{k2,1}$ captures a firm's existing network scale in the focal market, and $X_{k2,2}$ captures potential spatial spillover from adjacent tracts.

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 $X_{k2,1}$, and 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 D. These robustness checks only lead to marginally different point estimates, and conform to the main conclusion in our estimation and counterfactual analyses.

4.5.4 Summary statistics: market attributes

In Table 2, we compare the market attributes of the census tracts in the sample (2,582 in total) for our entry game and those of the remaining parts of the country (70,745 in total). 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 2, 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.

4.6 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_{k1}: tract attributes from 2016 ACS + the number of 4G-LTE incumbents in the focal census tract by the end of 2015;[40]

[40]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 1: 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.

Table 2: Census Tract Attributes

Variable	Definition	Markets to enter Mean	Markets to enter S.D.	Other markets Mean	Other markets S.D.
Pop (in 1,000'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 Year 2016 ($1,000'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.

- X_{k2}: includes two components – potential entrant k's 3G deployment in the focal census tract, $X_{k2,1}$, and its 4G-LTE deployment in neighboring tracts, $X_{k2,2}$;
- Z_k: instrumental variables for $X_{k2,1}$ and $X_{k2,2}$ (the past attributes of neighboring tracts in 2000);
- u_k: unobserved errors in the *ex post* payoffs (Equation (1));
- 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 (6))

In the specification above, we focus on 4G-LTE competition. That is, 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 instruments Z_k for these endogenous variables.

In a firm's decision rule (Equation (6)), 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. In contrast, V_k, 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 area surrounding the focal market.[41]

Valid instruments for X_{k2} need to be relevant for the "network investment" variables, to be excluded from the entry payoff function, and to be orthogonal to η_k and V_k. For each focal census tract, we use the demographics of its "neighbors" (i.e., other tracts in the same county) in 2000 as instruments for $X_{k2,1}$ and $X_{k2,2}$. Below we discuss why these instruments satisfy the three properties required. We go by the order of exclusion, relevance and exogeneity for the purpose of coherent argument.

First, exclusion. We argue that a potential entrant only evaluates the focal market's attributes, its 3G deployment and its own 4G-LTE deployment in neighboring tracts when evaluating its payoffs. One may worry that a potential entrant makes entry decisions on a much larger scale than a census tract, so the lagged demographic variables of neighboring tracts may enter the payoffs on the focal market directly. However, as explained earlier, our sample consists of tracts that had not been entered by most major providers as of 2016. These tracts were typically isolated, with the surrounding tracts already served by 4G-LTE before the start of our sampling period. This is shown in Table 1's summary statistics on the percentage of census blocks covered with 4G-LTE in the neighboring tracts. Therefore, we believe modeling the Big Four providers' post-2015 4G-LTE deployment decisions at the level of local markets serves as a first-order approximation that captures the firms' main strategic concerns.

Second, relevance. These lagged demographic variables of neighbors affect 4G-LTE deployment in the neighboring markets directly, and, therefore, are relevant for $X_{k2,2}$. They are relevant for $X_{k2,1}$ (the focal market's 3G deployment), not because that they directly enter a firm's decision to deploy 3G in the focal market, but because they affect a firm's 3G deployment in the neighboring markets. Note that 3G technology was actively deployed between 2000 and 2010; therefore, the 2000 Census's market attributes are very relevant to 3G deployment. This is a "spillover effect" through network investment, because 3G network nearby could lower the cost

[41]Each provider's spectrum holdings in 3G technology could even be considered a good instrument variable for $X_{k2,1}$ conceptually. Unfortunately, we could not locate such information in any database, as activities on secondary markets, such as leases, roaming agreements, and acquisitions, were frequent in the last two decades.

of 3G deployment in the focal market. Through the direct effects on the neighboring markets and the spillover effect to the focal market, these instruments are relevant to the two endogenous, “network investment" variables in our main equation.

Lastly, orthogonality. It is plausible that these lagged demographics on neighboring tracts are orthogonal to the unobserved factors determining deployments in focal and neighboring markets (η_k and V_k), once conditional on market-level observables. First, the correlation between 2000 Census's market attributes and 2016's market-level unobserved heterogeneity is weakened over 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) should be driven by the neighboring markets' attributes around 2016, after multiple generations of cellphone technology evolution and secondary market trading. Second, we believe that the detailed market-level attributes included in X_{k1} have already captured the spatial correlation across census tracts. Thus we abstract away from spatial correlation in the unobservables. That is, conditional on X_{k1}, the unobserved η_k and V_k, which capture firm- and market-specific unobserved heterogeneity, are not spatially correlated across census tracts. Lastly, we intentionally do not include “growth" variables such as population and income growth rates, because these variables are likely to contribute to current strategic considerations of entering a cluster of markets and therefore jeopardize the validity of our instrumental variables.

Our choice of neighboring markets' attributes as instruments originates from an insight in Fan (2013), which used the demographics in the market of a firm's competitors (excluding this firm's own market) as instruments. The competitors' non-overlapping markets are typically adjacent to a newspaper's market. In particular, Fan (2013) showed that the demographics of neighboring counties are not highly correlated (which implies her exogenous covariates and excluded instruments are not highly correlated). Our exogenous covariates and excluded instruments follow similar structures in our first-stage regressions. Many coefficients on neighboring markets' lagged demographics in Table 9 are statistically significant, suggesting these excluded instruments matter for our endogenous network deployment variables even with all the exogenous covariates of focal markets' attributes. 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 8.514 with a p-value less than 0.001; for $X_{k2,2}$, the F-statistic is 46.19 with p-value less than 0.001.

4.7 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" potential entrants and T-Mobile and Sprint as "weak" ones. 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 competitors. We hypothesize that "weak" potential entrants will suffer larger competition effects because they are less able to secure market shares facing competition. We present results treating all four firms as equal competitors in Appendix D. We adopt a specification in which the coefficients are not firm-specific (except for the aforementioned heterogeneous competition effects) in *ex post* payoffs. Thus, for simplicity, we suppress the index k in $X_{k2,1}$, $X_{k2,2}$, V_{k1}, V_{k2} when reporting our estimation and simulation results.

Table 3 presents estimation results from two models, with and without accounting for endogeneity in X_2 respectively. In the former model, we include all exogenous and instrumental variables and their squared terms (excluding dummy variables) in the first stage estimation (the first-stage results are reported in Appendix E). In the latter model, 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 larger effect on weak potential entrants as we hypothesized. The incumbent effect is also significantly negative. The incumbent competition effect is in fact much smaller that the expect "big four" competition effects (α_k). This is because most of these incumbents are small, local or regional providers that were not equal competitors with any of the Big Four firms. Population size contributes to 4G-LTE entry positively, but the percentages of seniors, Natives and Hispanics, as well as water coverage, act in the opposite direction. The percentage of labor force participation contributes to 4G-LTE entry negatively, which can be rationalized as labor force participation has no clean-cut relationship with the willingness to pay for mobile phone services. For example, if a census tract has a lot of low-end jobs that do not pay well, then higher labor force participation will not necessarily translate into greater willingness to subscribe to 4G-LTE. The percentage of rural population, likewise, has an ambiguous relationship with such willingness to pay. For example, residents who live in a rural area may be willing to spend on better mobile coverage to stay connected.

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.[42] 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<3.757).

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.[43] Hence, there is 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>1.618).

Table 3 also shows that ignoring endogeneity in 3G and neighboring 4G-LTE

[42]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.

[43]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).

deployment only leads to slight exaggeration of competition effects (α_k). The difference between these estimates is small relative to the size of marginal effects by the network deployments (γ_k), and does not appear statistically significant given the standard errors. This may be related to our hypothesis that the source for endogeneity, i.e., the correlation between u_j and V_j, is largely due to the latent, firm-specific spectrum holdings. Thus, not accounting for such a source of endogeneity in estimation would leave the structural errors more directly and strongly correlated with the individual firm's deployment (whose coefficient is γ_k) than with the other common features that affect the equilibrium beliefs (whose coefficient is α_k).

4.8 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 3 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 an integrated T-Mobile and Sprint network (henceforth referred to as a "New T-Mobile").[44] 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 initially 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.[45]

We use the estimated coefficients in Table 3 to simulate the local market entry decisions of Verizon, AT&T and New T-Mobile (and DISH in the third scenario). For

[44] 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 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.

[45] 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 3: 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 1,000's)	0.078***	0.020	0.081***	0.019
% Female	0.437	0.761	0.417	0.758
% Senior	-1.035**	0.414	-1.110***	0.425
% White	-0.233	0.513	-0.170	0.515
% Black	-0.351	0.526	-0.367	0.535
% Native	-1.396***	0.536	-1.611***	0.518
% Asian	0.870	1.310	1.290	1.316
% Hispanic	-0.469	0.339	-0.185	0.308
% College	-0.150	0.403	-0.063	0.417
% Labor force	-0.943***	0.322	-0.627**	0.314
% Work home	0.289	0.522	0.768	0.588
% Long comm.	0.069	0.307	-0.508*	0.297
HH income	-0.002	0.003	-0.006**	0.003
HH size	-0.042	0.026	-0.059**	0.026
Pop density	-0.131	0.089	-0.229***	0.084
% Rural	0.223*	0.126	0.188**	0.092
Mostly water	-1.850***	0.550	-2.418***	0.545
# Incumbents	-0.158***	0.043	-0.225***	0.042
Intercept	0.911	0.784	0.871	0.757
Network Investment Effects				
$X_{2,1}$	3.757***	0.790	2.024***	0.178
$X_{2,2}$	1.618***	0.323	3.042***	0.096
V_1	-1.738**	0.784	–	–
V_2	1.536***	0.335	–	–
Expected Competition Effects				
Strong potential entrant	-1.070***	0.090	-1.084***	0.085
Weak potential entrant	-1.105***	0.061	-1.133***	0.064

Notes: The results are based on 6,244 potential entrants for 2,582 census tracts (each with at least two potential big-4 entrants). Standard errors are obtained by resampling the markets with replacement for 1,000 times. Asterisks indicate the statistical significance at the 1% (***), 5% (**), and 10% (*) levels. The estimated competitive effects satisfy the moderate interaction condition in Assumption 2 on 95.7% of markets in the sample.

comparison, in each scenario, we simulate two sets of outcomes, one with and one without accounting for endogenous X_2. Panel A in Table 4 presents the simulated market entries across different scenarios; Panel B in Table 4 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 4 are simulation results under the three scenarios, using structural estimates that account for endogeneity (column (1) of Table 3). Comparing column (1) to column (2), we can see that the T-Mobile and Sprint merger reduces the number of total entry occurrences from 2,898 to 2,264, which is a 21.88% reduction. This leads to a large increase (23.15%) in the underserved population. There are two explanations for such a reduction: First, there are fewer potential entrants in the markets after the merger. Second, the New T-Mobile resulting from the merger is a strong competitor with an integrated network from Sprint and T-Mobile and, therefore, the entry probability of the merged firm is higher than that of T-Mobile or Sprint alone. That also impacts how the merger deters entry by the other competitors. These two effects can also explain the simulated increase in the number of markets with no Big-4 entrant under the T-Mobile/Sprint merger in column (2). On the one hand, the post-merger New T-Mobile may well benefit from an economy of scale due to the combined network investment, and consequently has stronger profit incentives for entry; on the other hand, our simulation indicates that such an increase in the entry likelihood of the single, merged entity is not substantive enough to off-set the negative impact on entry occurrence due to fewer potential entrants and the equilibrium responses by the other Big-4 competitors (AT&T and Verizon). 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.

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 4 reports that the number of entry occurrences is 2,940, substantially higher than in the scenario without the divestiture required by the DOJ. This practically restores the level of market entries in the baseline scenario without the merger. On most dimensions of entry outcomes and population covered, the DISH remedy seems to be effective. Despite the entry occurrences under the DISH remedy being even higher than the baseline (2,940 vs. 2,898), the total population

Table 4: 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	522	633	516	524	578	521
n entrants = 1	1,277	1,641	1,244	1,277	1,677	1,234
n entrants = 2	730	300	771	730	319	770
n entrants = 3	50	8	51	49	8	53
n entrants = 4	2	–	–	3	0	5
Total # entry occurrences	2,898	2,264	2,940	2,893	2,338	2,951
by AT&T	716	727	714	716	740	717
by Verizon	157	171	150	155	186	152
by T-Mobile/New T-Mobile	1,201	1,366	1,262	1,194	1,412	1,258
by Sprint/DISH Network	824	–	814	827	–	824
Panel B: Population (in 1,000's) under-served (# incumbents in 2018 ≤ 1)						
Underserved population	Baseline	4 to 3	DISH	Baseline	4 to 3	DISH
Total population	216	266	223	226	263	224
Minority population	122	135	124	124	132	123
Rural population	162	196	165	167	196	166

Notes: For each market, we make 1,000 random draws of the error vector, and use the structural estimates from Table 3 to generate 1,000 counterfactual predictions. The table reports average predictions (rounded to the nearest integer). The above results are based on 2,582 census tracts, each with at least one potential entrant from the Big Four. 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.

under-served would increase by 3.24% from column (1) to column (3). We checked the 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 the markets entered is a strategic, equilibrium response due to the different configurations of network investment and competitors of different statuses in the markets.

Columns (4) to (6) of Table 4 report simulation results using estimates that do not

Table 5: 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
Total # entry occurrences	-21.88%	-19.18%	1.45%	2.00%
# markets with 0 entrant	21.26%	10.31%	-1.15%	-0.57%
# markets with 1 entrant	28.50%	31.32%	-2.58%	-3.37%
Population (in 1,000's) under-served (# incumbents in 2018 ≤ 1)				
Total population	23.15%	16.37%	3.24%	-0.88%
Minority population	10.66%	6.45%	1.64%	-0.81%
Rural population	20.99%	17.37%	1.85%	-0.60%

Notes: The above results are calculated using Table 4 results.

account for endogeneity in 3G and neighboring 4G-LTE deployment (the first column in panel (2) of Table 3). A comparison between column (1) and column (4) suggests the impact of ignoring endogeneity in X_2 on simulated entries under the baseline scenario is negligible. The predicted changes in market entry under the second (4-to-3 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 5 to compare the magnitudes of merger effects and remedy effects under these two models.

As Table 5 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 21.26% with the 4-to-3 merger under the endogenous X_2 model but only 10.31% if the endogeneity in X_2 is ignored in estimation and simulation. Underestimation of a similar magnitude also happens for population affected by reduced entry. For example, the merger will lead to a 23.15% increase in the population underserved and a 20.99% increase in the rural population underserved under the endogenous X_2 model but only 16.37% and 17.37% 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 are more optimistic than those predicted under the model with endogenous X_2. The model with exogenous X_2 slightly overestimates the percentage increase in entry occurrence under the DISH remedy. It also suggests the DISH remedy would even lead to small reductions in the underserved population, while the model that accounts for endogenous X_2 indicates the remedy will actually lead to non-trivial increases in the under-served population. The key message is that researchers would paint a much rosier picture of the consequences due to the merger and the impact of the DISH remedy if they were to use a model that does not account for X_2's endogeneity. Relying on such a biased prediction, policymakers would lean more toward approving the proposed merger.

5 Conclusion

The literature on discrete choice games is silent on how to deal with endogenous covariates, which is often the focus of empirical exercises. An example of potentially endogenous covariates in firm profits from market entry is the continuous measure of airport presence in Berry (1992). Another example is a grocery store's distance to a supercenter in Grieco (2014), as supercenter locations are endogenous outcomes of a related game at a grander level and possibly correlate with the unobserved heterogeneity of rural markets to be entered by smaller stores. In these studies, the potentially endogenous variables are treated as exogenous. The empirical literature so far has not provided researchers with means to accommodate both the effect of competition (or social interaction) and that of the endogenous covariates in a discrete game setting.

We fill in this gap, proposing a new method for estimating discrete games with incomplete information in the presence of 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 in our sample 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.[46]

Individuals and firms weigh their decisions based on all available information in strategy settings, but not all of it is captured by observed covariates in empirical models. Some observed covariates are correlated with the player- and game-level unobserved heterogeneity. There are many analogies in applied microeconomics, in which it is essential that researchers incorporate endogenous variables to study a key decision. In labor economics, a worker's career choice depends on her accumulation of human capital. In industrial organization, a pharmaceutical company's patenting decision depends on its stock and flow of innovative activities; a cable TV network's decision to discontinue a TV series depends on the series' rating and viewership. We hope our method will provide a useful tool to identify multiple parameters of interest and, in turn, improve researchers' ability to diagnose, predict and recommend policy remedies.

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.

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

Arcidiacono, P. and R. A. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica 79*(6), 1823–1867.

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. and P. Reiss (2007). Empirical models of entry and market structure. *Handbook of industrial organization 3*, 1845–1886.

Berry, S. T. (1992). Estimation of a model of entry in the airline industry. *Econometrica 60*(4), 889–917.

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., Y. Sun, and F. Verboven (2021). Market entry, fighting brands, and tacit collusion: Evidence from the french mobile telecommunications market. *American Economic Review 111*(11), 3459–99.

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 129*(11), 2995–3038.

Collard-Wexler, A. (2013). Demand fluctuations in the ready-mix concrete industry. *Econometrica 81*(3), 1003–1037.

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.

Dunne, T., S. D. Klimek, M. J. Roberts, and D. Y. Xu (2013). Entry, exit, and the determinants of market structure. *The RAND Journal of Economics 44*(3), 462–487.

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.

Gu, X., H. Li, Z. Lin, and X. Tang (2022). Peer effects with sample selection. *Working paper*.

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.

Hotz, V. J. and R. A. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies 60*(3), 497–529.

Igami, M. and N. Yang (2016). Unobserved heterogeneity in dynamic games: Cannibalization and preemptive entry of hamburger chains in canada. *Quantitative Economics 7*(2), 483–521.

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 (2022). Adjusting for peer-influence in propensity scoring when estimating treatment effects. *Working paper*.

Jia, P. (2008). What happens when wal-mart comes to town: An empirical analysis of the discount retailing industry. *Econometrica 76*(6), 1263–1316.

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. and X. Tang (2022). Solving reflection problems in social interactions models with endogeneity. *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. (2022). Strategic interactions in mobile network investment with a new entrant and unobserved heterogeneity. *International Journal of Industrial Organization 82*, 102829.

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 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 $\theta \in \Theta$, $P \in \mathcal{P}$,

$$L_n(\theta, P; \widehat{\Pi}) - L_n(\theta, P; \Pi_0) = \nabla_\Pi L_n(\theta, P; \Pi^+)(\widehat{\Pi} - \Pi_0), \tag{12}$$

where Π^+ denotes an intermediate value between $\widehat{\Pi}$ and Π_0. By (12) 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

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

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

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

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

$$\tau \equiv \min_{(\theta, P) \in B_\epsilon^c(\theta_0, P^0) \cap \mathcal{E}} T(\theta, P; \Pi_0) > 0. \tag{13}$$

Define the event

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

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

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

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

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

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

$$\begin{aligned} A_n \quad & \Rightarrow \quad \left\{ L_0\left(\theta_n^*, P_n^*; \Pi_0\right) > L_0\left(\theta, P_n^*; \Pi_0\right) - \tau, \forall \theta \in \Theta \right\}, \\ & \Rightarrow \quad \left\{ L_0\left(\theta_n^*, P_n^*; \Pi_0\right) > \max_{\theta \in \Theta} L_0\left(\theta, P_n^*; \Pi_0\right) - \tau \right\}, \\ & \Rightarrow \quad \left\{ \tau > T\left(\theta_n^*, P_n^*; \Pi_0\right) \right\}, \\ & \Rightarrow \quad \left\{ \min_{(\theta, P) \in B_\epsilon^c(\theta_0; P^0) \cap \mathcal{E}} T(\theta, P; \Pi_0) > T\left(\theta_n^*, P_n^*; \Pi_0\right) \right\} \text{ by (13)}, \\ & \Rightarrow \quad \left\{ (\theta_n^*, P_n^*) \in B_\epsilon(\theta_0, P^0) \right\}. \end{aligned}$$

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,

$$\Pr\left((\theta_n^*, P_n^*) \in B_\epsilon(\theta_0; P^0)\right) \to 1. \tag{14}$$

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,

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

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.

$\square$

Appendix B Monte Carlo Evidence

In this appendix, 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.

B.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 6: 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 7: 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

B.2 Heterogeneous competitive effects

We now 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 $\alpha_3, \alpha_4 = \alpha_W$.[47]

[47] 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.

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).

Appendix C 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.[48] In the 1990s, 2G started the use of digital transmission 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 D Robustness of Table 3

We check the robustness of our Table 3'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 3, 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 3, 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 3, 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 3.

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.

[48] 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.

Appendix E First-stage Results

Table 9 reports the first-stage regression results for the endogenous network investment model. In the reported results in this paper, we adopt a specification with covariates of X_1, Z, and their squared terms in the first stage regressions of $X_{2,1}$ and $X_{2,2}$. We reported the first-stage results in Spec. (2), corresponding to columns (2) and (4) in Table 9. In Spec. (2), the adjusted R^2's for $X_{2,1}$ and $X_{2,2}$ regressions are 0.077 and 0.333, respectively. In Spec. (1), we report the first-stage results in which we include only (X_1, Z) — but not the squared terms — as the regressors.

Table 8: Robustness of Table 3

	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)	0.078***	0.020	0.071***	0.015	0.051***	0.015
% Female	0.436	0.758	0.711	0.520	0.399	0.520
% Senior	-1.046**	0.413	-0.954***	0.338	-0.772**	0.338
% White	-0.224	0.513	-0.235	0.395	-0.471	0.395
% Black	-0.351	0.527	-0.359	0.413	-0.728*	0.413
% Native	-1.391***	0.537	-1.626***	0.398	-1.045***	0.398
% Asian	0.877	1.314	1.181	1.036	0.310	1.036
% Hispanic	-0.468	0.340	0.042	0.232	-0.467**	0.232
% College	-0.155	0.402	0.103	0.320	0.180	0.320
% Labor force	-0.943***	0.321	-0.846***	0.243	-1.108***	0.243
% Work home	0.277	0.521	0.733*	0.430	0.051	0.430
% Long commute	0.070	0.307	-0.144	0.217	0.439**	0.217
HH income	-0.002	0.003	-0.006**	0.002	0.001	0.002
HH size	-0.042	0.026	-0.043**	0.019	-0.013	0.019
Pop density	-0.132	0.089	-0.192***	0.070	-0.051	0.070
% Rural	0.221*	0.126	0.021	0.074	00.116	0.074
Mostly Water	-1.858***	0.548	-2.438***	0.410	-1.383***	0.410
# Incumbents	-0.162***	0.043	-0.188***	0.034	-0.007	0.034
Intercept	0.926	0.784	1.256**	0.540	0.670	0.540
Network Investment Effects						
$X_{2,1}$	3.743***	0.788	1.261***	0.235	3.705***	0.235
$X_{2,2}$	1.610***	0.322	1.958***	0.119	0.254**	0.119
V_1	-1.731**	0.783	0.704***	0.183	1.103***	0.183
V_2	1.531***	0.335	1.114***	0.204	0.472**	0.204
Competition Effects						
Any P.E.	-1.099***	0.062	-	-	-	-
Strong P.E.	-	-	-1.012***	0.084	-0.436***	0.084
Weak P.E.	-	-	-1.086***	0.064	-0.352***	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.

Table 9: First Stage Regressions with Different Specifications

	$X_{2,1}$		$X_{2,2}$	
Variable	Spec.(1)	Spec.(2)	Spec.(1)	Spec.(2)
Focal Pop (in 000's)	0.001	0.024**	-0.002	0.009
Focal % Female	0.075	0.086	0.157	0.081
Focal % Senior	0.009	0.269	0.126**	0.171
Focal % White	0.105	0.122	-0.394**	0.340
Focal % Black	0.047	0.054	-0.192	-0.127
Focal % Native	0.010	0.012	-0.179*	-0.160
Focal % Asian	-0.231	-0.739**	-0.433**	-0.229
Focal % Hispanic	0.160***	0.277**	-0.091	-0.120
Focal % College	0.109	0.222	0.123*	-0.422**
Focal % Labor force	-0.140**	-0.578**	-0.341***	-0.883***
Focal % Work home	-0.014	0.084	-0.327***	-0.744***
Focal % Long comm.	-0.014	-0.198	0.342***	0.396***
Focal HH income	-0.001**	-0.002**	0.000	0.001
Focal HH size	0.000	0.016	0.011***	0.016
Focal Pop density	-0.005	0.001	-0.011	0.002
Focal % Rural	-0.098***	0.003	-0.080***	-0.111
Focal Mostly water	-0.235***	-0.184***	0.066	0.083
Focal # Incumbents	-0.017***	-0.076***	0.021***	0.026
Excluded Instruments				
Lagged Pop (in 000's)	0.003	0.049*	0.030***	0.023
Lagged % Female	0.203	4.858	0.416	1.851
Lagged % Senior	0.265**	0.164	0.212*	-0.102
Lagged % White	0.544***	0.257	0.135	0.394
Lagged % Black	-0.392**	-0.883**	-0.061	-0.983***
Lagged % Native	-0.556***	-0.731**	-0.179	-0.844**
Lagged % Asian	-0.338	0.998	-0.331	-0.293
Lagged % Hispanic	-0.050	0.043	0.109	-0.502**
Lagged % College	0.069	-0.294	0.143	-0.825***
Lagged % Labor force	0.089	-0.062	-0.177*	3.387***
Lagged % Work home	0.200	-0.018	-0.225	-1.153**
Lagged % Long comm.	0.078	0.223	-0.022	0.447**
Lagged HH income	0.001**	0.010***	0.003***	0.025***
Lagged HH size	0.011	0.030	0.014	0.046**
Lagged Pop density	-0.006	0.002	0.014	0.054***
Lagged % Rural	0.007	0.204**	-0.151***	0.277***
Lagged Mostly water	-0.496***	-0.780***	-0.591***	-1.023***
Intercept	0.408*	-1.385	0.329	-1.597
Including squared terms	No	Yes	No	Yes
# of Regressors	35	69	35	69
Adjusted R^2	0.068	0.077	0.306	0.333

Notes: To save space, we only report estimates for (X_1, Z) in specification (2). All excluded instruments are lagged market attributes in neighboring markets.